

Nonnegative Inverse Eigenvalue Problems with Partial Eigendata

Zheng-Jian Bai* Stefano Serra-Capizzano[†] Zhi Zhao[‡]

June 25, 2011

Abstract

In this paper we consider the inverse problem of constructing an n -by- n real nonnegative matrix A from the prescribed partial eigendata. We first give the solvability conditions for the inverse problem without the nonnegative constraint and then discuss the associated best approximation problem. To find a nonnegative solution, we reformulate the inverse problem as a monotone complementarity problem and propose a nonsmooth Newton-type method for solving its equivalent nonsmooth equation. Under some mild assumptions, the global and quadratic convergence of our method is established. We also apply our method to the symmetric nonnegative inverse problem and to the cases of prescribed lower bounds and of prescribed entries. Numerical tests demonstrate the efficiency of the proposed method and support our theoretical findings.

Keywords. Nonnegative matrix, inverse problem, monotone complementarity problem, generalized Newton's method, semismooth function.

AMS subject classifications. 49J52, 49M15, 65F18, 90C33

1 Introduction

Inverse eigenvalue problems (IEPs) arise in a wide variety of applications such as structural dynamics, control design, system identification, seismic tomography, remote sensing, geophysics, particle physics, and circuit theory, etc. For the applications, mathematical properties, and algorithmic aspects of general IEPs, we may refer to [8, 9, 12, 22, 24, 47] and references therein.

An m -by- n matrix $M \geq 0$ ($M > 0$, respectively) is called nonnegative (strictly positive, respectively) if $M_{ij} \geq 0$ ($M_{ij} > 0$, respectively) for all $i = 1, \dots, m$ and $j = 1, \dots, n$. Nonnegative matrices play an important role in many applications such as game theory, Markov chain,

*School of Mathematical Sciences, Xiamen University, Xiamen 361005, People's Republic of China, Dipartimento di Fisica e Matematica, Università dell'Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy (zjbai@xmu.edu.cn). The research of this author was partially supported by the Natural Science Foundation of Fujian Province of China for Distinguished Young Scholars (No. 2010J06002), NCET, and Internationalization Grant of U. Insubria 2008, 2009.

[†]Dipartimento di Fisica e Matematica, Università dell'Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy (stefano.serrac@uninsubria.it). The work of this author was partially supported by MIUR (No. 20083KLJEZ).

[‡]School of Mathematical Sciences, Xiamen University, Xiamen 361005, People's Republic of China (zhaozhi231@163.com).

probabilistic algorithms, numerical analysis, discrete distributions, categorical data, group theory, matrix scaling, and economics, etc. One may refer to [1, 2, 29, 40] for the applications and mathematical properties of nonnegative matrices. The nonnegative inverse eigenvalue problem has got much attention since 1940s (see the survey papers [8, 15] and references therein). Most of the works determine necessary and sufficient conditions such that the given complete set of complex numbers is the spectrum of a nonnegative matrix. Recently, despite theoretical incompleteness of the nonnegative inverse problem, there are few numerical algorithms developed for computational purpose, including the isospectral flow method [5, 6, 7, 10] and the alternating projection method [31].

In this paper, we consider the inverse problem of constructing a real nonnegative matrix from the given partial eigendata. The canonical problem can be stated as follows.

NIEP. Construct a nontrivial n -by- n nonnegative matrix A from a set of measured partial eigendata $\{(\lambda_k, x_k)\}_{k=1}^p$ ($p \leq n$).

In practice, the entries of a nonnegative matrix denote the distributed physical parameters such as mass, length, elasticity, inductance, capacitance, and so on. Moreover, an 'a priori' analytical nonnegative matrix A_a can be obtained by using the finite element techniques. However, the predicted dynamical behaviors (eigenvalues and eigenvectors) of the analytical nonnegative matrix A_a often disagree with the experimental eigendata [22]. The inverse problem aims to reconstruct a physically feasible nonnegative matrix from the measured eigendata.

In this paper, we first give the solvability conditions for the NIEP, without the nonnegative constraint, and then we study the corresponding best approximation problem with respect to the analytical nonnegative matrix A_a . To find a physical solution to the NIEP, we reformulate the NIEP as a monotone complementary problem (MCP) and we propose a nonsmooth Newton-type method for solving its equivalent nonsmooth equation. In fact, we are motivated by recent developments of Newton-type methods for MCPs [13, 16, 18, 19, 23, 25, 26, 48]. Under some mild conditions, we show that the proposed method converges globally and quadratically. We also use our method to the symmetric nonnegative inverse problem and to the cases of prescribed lower bounds and of prescribed entries. Numerical tests are also reported, in order to illustrate the efficiency of our method and to confirm our theoretical results.

Throughout the paper, we use the following notations. The symbols M^T, M^H , and M^+ denote the transpose and the conjugate transpose, and the Moore-Penrose inverse of a matrix M , respectively. I is the identity matrix of appropriate dimension. We denote by $\|\cdot\|$ the Euclidean vector norm and the Frobenius matrix norm, and by V^* the adjoint of an operator V . The symbols $\mathbb{R}_+^{m \times n}$ and $\mathbb{R}_{++}^{m \times n}$ stand for the nonnegative orthant and the strictly positive orthant of $\mathbb{R}^{m \times n}$, respectively. Given $\mathcal{N} := \{(i, j) \mid i, j = 1, \dots, n\}$, the sets of indices $\mathcal{I}, \mathcal{J} \subseteq \mathcal{N}$ are such that $\mathcal{J} = \mathcal{N} \setminus \mathcal{I}$. Let $|\mathcal{I}|$ be the cardinality of the index set \mathcal{I} . For a matrix $M \in \mathbb{R}^{n \times n}$, $M_{\mathcal{I}}$ is the column vector with entries M_{ij} for all $(i, j) \in \mathcal{I}$. Define the linear operator $P : \mathbb{R}^{|\mathcal{I}|} \rightarrow \mathbb{R}^{n \times n}$ by

$$P_{ij}(M_{\mathcal{I}}) := \begin{cases} M_{ij}, & \text{if } (i, j) \in \mathcal{I}, \\ 0, & \text{otherwise.} \end{cases}$$

The paper is organized as follows. In Section 2 we study the solvability of the NIEP by neglecting the nonnegative requirement and we discuss the associated best approximation problem with respect to an 'a priori' analytical nonnegative matrix. In Section 3 we reformulate the

NIEP as a MCP and we propose a Newton-type method for solving its equivalent nonsmooth equation. Under some mild conditions, the global and quadratical convergence of our method is established. In Section 4 we discuss the application of our method to specific important cases. In Section 5 we report some numerical results to illustrate the efficiency of the proposed method.

2 Solvability and Best Approximation

In this section, we give a sufficient and necessary condition for the solvability of the NIEP without the nonnegative constraint and then we discuss the associated best approximation problem with respect to an 'a priori' analytical nonnegative matrix A_a . We note that the complex eigenvalues and eigenvectors of a nonnegative matrix $A \in \mathbb{R}^{n \times n}$ appear in complex conjugate pairs. If $a \pm b\sqrt{-1}$ and $x \pm y\sqrt{-1}$ are complex conjugate eigenpairs of A , where $a, b \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$, then we have

$$Ax = ax - by \quad \text{and} \quad Ay = ay + bx,$$

or

$$A[x, y] = [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Therefore, without loss of generality, we can assume

$$\Lambda = \text{diag}(\lambda_1^{[2]}, \dots, \lambda_s^{[2]}, \lambda_{2s+1}, \dots, \lambda_p) \in \mathbb{R}^{p \times p}$$

and

$$X = [x_{1R}, x_{1I}, \dots, x_{sR}, x_{sI}, x_{2s+1}, \dots, x_p] \in \mathbb{R}^{n \times p},$$

where

$$\lambda_i^{[2]} = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}, \quad \lambda_{2i-1, 2i} = a_i \pm b_i\sqrt{-1}, \quad x_{2i-1, 2i} = x_{iR} + x_{iI}\sqrt{-1}, \quad a_i, b_i \in \mathbb{R}, \quad x_{iR}, x_{iI} \in \mathbb{R}^n,$$

for all $1 \leq i \leq s$. By neglecting the nonnegative constraint, the NIEP reduces to the following problem:

Problem 1. Construct a nontrivial n -by- n real matrix A from the measured eigendata $(\Lambda, X) \in \mathbb{R}^{p \times p} \times \mathbb{R}^{n \times p}$.

To study the solvability of Problem 1, we need the following preliminary lemma.

Lemma 2.1 [45, Lemma 1.3] *Given $B_1 \in \mathbb{C}^{q \times m}$, $B_2 \in \mathbb{C}^{n \times t}$, $B_3 \in \mathbb{C}^{q \times t}$, $Y_a \in \mathbb{C}^{m \times n}$. Define*

$$\mathcal{E} := \{Y \in \mathbb{C}^{m \times n} : B_1 Y B_2 = B_3\}.$$

Then $\mathcal{E} \neq \emptyset$ if and only if B_1, B_2, B_3 satisfy

$$B_1 B_1^+ B_3 B_2^+ B_2 = B_3.$$

In the case of $\mathcal{E} \neq \emptyset$, any $Y \in \mathcal{E}$ can be expressed as

$$Y = B_1^+ B_3 B_2^+ + G - B_1^+ B_1 G B_2 B_2^+,$$

where $G \in \mathbb{C}^{m \times n}$. Moreover, there is a unique matrix $\bar{Y} \in \mathcal{E}$ given by

$$\bar{Y} = B_1^+ B_3 B_2^+ + Y_a - A_1^+ A_1 Y_a A_2 A_2^+$$

such that for any unitarily invariant norm $\|\cdot\|$,

$$\|\bar{Y} - Y_a\| = \min_{Y \in \mathcal{E}} \|Y - Y_a\|.$$

Based on Lemma 2.1, we easily derive the following result on the solvability of Problem 1 and its approximation.

Theorem 2.2 *Problem 1 has a solution if and only if $X\Lambda X^+ X = X\Lambda$. In this case, the general solution to Problem 1 is given by*

$$A = X\Lambda X^+ + G(I - XX^+),$$

where $G \in \mathbb{R}^{n \times n}$ is arbitrary. Moreover, for an 'a priori' nonnegative matrix $A_a \in \mathbb{R}^{n \times n}$, there exists a unique solution to Problem 1 given by

$$\bar{A} = X\Lambda X^+ + A_a(I - XX^+)$$

such that $\|\bar{A} - A_a\| = \min_{A \in \mathcal{S}} \|A - A_a\|$, where \mathcal{S} is the solution set of Problem 1.

Remark 2.3 *We note from Theorem 2.2 that Problem 1 is solvable if X is full column rank. Furthermore, we can find a unique solution in the solution set of Problem 1, which is nearest to a fixed a priori nonnegative matrix and satisfies the given eigendata though it may not be physically feasible.*

For the sake of clarity, we give two numerical examples.

Example 2.4 *Let $n = 6$. We randomly generate an n -by- n nonnegative matrix \hat{A} as follows:*

$$\hat{A} = \begin{bmatrix} 0.8270 & 0.3112 & 0.8260 & 0.9632 & 0.5067 & 0.1420 \\ 0.5522 & 1.0324 & 0.8392 & 0.3307 & 0.7635 & 0.6059 \\ 1.0387 & 0.4184 & 0.9698 & 0.4000 & 1.0901 & 0.4353 \\ 0.3360 & 0.4230 & 0.7811 & 0.9965 & 0.8516 & 0.6115 \\ 0.1277 & 0.5167 & 0.6465 & 0.8481 & 0.7110 & 0.5592 \\ 0.2316 & 0.7494 & 1.0024 & 0.8008 & 0.8709 & 0.8055 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\lambda_1 = 3.9752$, $\lambda_{2,3} = 0.6941 \pm 0.2340\sqrt{-1}$, $\lambda_4 = -0.2290$, $\lambda_{5,6} = 0.1039 \pm 0.0572\sqrt{-1}$. We use the first three eigenvalues $\{\lambda_i\}_{i=1}^3$ and associated eigenvectors $\{x_i\}_{i=1}^3$ as prescribed eigendata.

For Example 2.4, we can easily check that the given eigendata satisfy the solvability condition in Theorem 2.2. Therefore, Problem 1 is solvable and the minimum norm solution is given by

$$A_{\min} = \begin{bmatrix} 0.8914 & 0.1511 & 0.8235 & 0.7375 & 0.4717 & 0.4671 \\ 0.5589 & 1.0095 & 0.7783 & 0.4834 & 0.4562 & 0.7772 \\ 0.9454 & 0.4153 & 0.9468 & 0.7767 & 0.5343 & 0.6213 \\ 0.5413 & 0.3943 & 0.5753 & 0.9115 & 0.7712 & 0.8074 \\ 0.3489 & 0.4857 & 0.4337 & 0.7302 & 0.6675 & 0.7523 \\ 0.5193 & 0.7965 & 0.6736 & 0.8095 & 0.7427 & 0.9366 \end{bmatrix}.$$

Suppose that an 'a priori' analytic nonnegative matrix A_a takes the form of

$$A_a = \begin{bmatrix} 0.7919 & 0.3850 & 0.8504 & 1.0241 & 0.4175 & 0.1737 \\ 0.5845 & 0.9893 & 0.8272 & 0.3646 & 0.7035 & 0.6031 \\ 1.0657 & 0.4856 & 1.0630 & 0.3690 & 1.0876 & 0.5047 \\ 0.3660 & 0.4342 & 0.7591 & 0.9863 & 0.8870 & 0.6477 \\ 0.0571 & 0.4183 & 0.6005 & 0.8866 & 0.7658 & 0.5215 \\ 0.1669 & 0.6504 & 1.0721 & 0.8716 & 0.8504 & 0.8913 \end{bmatrix}.$$

Then the best approximation is given by

$$\bar{A} = \begin{bmatrix} 0.7966 & 0.3178 & 0.8349 & 1.0390 & 0.4199 & 0.1502 \\ 0.5610 & 1.0346 & 0.8192 & 0.3603 & 0.7130 & 0.6281 \\ 1.0219 & 0.3927 & 0.9970 & 0.3733 & 1.0895 & 0.4703 \\ 0.3665 & 0.4201 & 0.7559 & 0.9686 & 0.8680 & 0.6264 \\ 0.1044 & 0.5193 & 0.6727 & 0.8492 & 0.7328 & 0.5328 \\ 0.1579 & 0.7029 & 1.0806 & 0.8120 & 0.8047 & 0.8726 \end{bmatrix}.$$

We observe that a physically realizable solution is obtained.

Example 2.5 Let \hat{A} be a random tridiagonal nonnegative matrix of order $n = 6$, where

$$\hat{A} = \begin{bmatrix} 4.7270 & 0.2055 & 0 & 0 & 0 & 0 \\ 0.4246 & 4.4522 & 0.2058 & 0 & 0 & 0 \\ 0 & 0.7618 & 4.9387 & 0.8847 & 0 & 0 \\ 0 & 0 & 0.7349 & 4.2360 & 0.2647 & 0 \\ 0 & 0 & 0 & 0.7497 & 4.0277 & 1.0682 \\ 0 & 0 & 0 & 0 & 0.2471 & 4.1316 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\{5.6126, 4.8973, 4.6442, 4.2472, 3.8061, 3.3058\}$. We use the first eigenvalues $\{\lambda_i\}_{i=1}^2$ and associated eigenvectors $\{x_i\}_{i=1}^2$ as prescribed eigendata.

For Example 2.5, the condition of Theorem 2.2 holds and thus Problem 1 is solvable. In particular, the minimum norm solution is given by

$$A_{\min} = \begin{bmatrix} 1.7782 & 1.6477 & 0.6440 & -0.5508 & -1.3782 & -0.4997 \\ 1.6687 & 1.6332 & 1.1661 & -0.1692 & -1.0924 & -0.4315 \\ 0.7795 & 1.2841 & 3.9140 & 2.0059 & 0.6949 & 0.0229 \\ -0.4669 & -0.0851 & 2.0783 & 1.5353 & 1.1657 & 0.2809 \\ -1.3297 & -1.0312 & 0.8174 & 1.2157 & 1.4952 & 0.4602 \\ -0.4906 & -0.4172 & 0.0643 & 0.3017 & 0.4668 & 0.1540 \end{bmatrix}.$$

Suppose that an 'a priori' analytic symmetric tridiagonal oscillatory matrix A_a takes the following form

$$A_a = \begin{bmatrix} 4.6799 & 0.3053 & 0 & 0 & 0 & 0 \\ 0.4981 & 4.3715 & 0.1122 & 0 & 0 & 0 \\ 0 & 0.8478 & 4.9398 & 0.9758 & 0 & 0 \\ 0 & 0 & 0.6451 & 4.2334 & 0.2571 & 0 \\ 0 & 0 & 0 & 0.8015 & 3.9347 & 1.0967 \\ 0 & 0 & 0 & 0 & 0.1885 & 4.2020 \end{bmatrix}.$$

Then the best approximation is given by

$$\bar{A} = \begin{bmatrix} 4.6636 & 0.2887 & -0.0158 & -0.0010 & 0.0092 & 0.0039 \\ 0.5104 & 4.3935 & 0.1854 & 0.0388 & 0.0151 & 0.0011 \\ -0.0184 & 0.8236 & 4.8874 & 0.9532 & -0.0021 & 0.0021 \\ 0.0089 & 0.0177 & 0.7094 & 4.2684 & 0.2720 & 0.0016 \\ -0.0174 & -0.0161 & -0.0065 & 0.8067 & 3.9480 & 1.1016 \\ -0.0093 & -0.0070 & 0.0073 & 0.0095 & 0.1996 & 4.2054 \end{bmatrix}.$$

We see that the best approximation is not physically realizable, but the tridiagonal entries are all positive and the off-tridiagonal entries are relatively small.

3 The Nonnegative Inverse Eigenvalue Problem with Partial Eigendata

In this section, we reformulate the NIEP as a MCP and propose a generalized Newton-type method for solving an equivalent nonsmooth equation to the MCP. In the following subsections, we first review some preliminary definitions and basic results for the nonlinear nonnegative programming which are used later in this paper. Then, we present a nonsmooth Newton-type method for solving the NIEP and give the convergence analysis.

3.1 Preliminaries

Let \mathcal{X} , \mathcal{Y} , \mathcal{Z} be three finite dimensional real vector spaces, each equipped with a scalar inner product $\langle \cdot, \cdot \rangle$ and the related induced norm $\|\cdot\|$. Sparked by the concepts of strong second-order sufficient condition and constraint nondegeneracy in the nonlinear semidefinite programming problem [43] and by the differential properties of the metric projector over the semidefinite cone [32], in this subsection, we briefly discuss some analogous definitions and basic properties for the following nonlinear nonnegative programming (NLNNP)

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0, \\ & g(x) \in \mathbb{R}_+^{n \times n}, \end{aligned}$$

where $f : \mathcal{X} \rightarrow \mathbb{R}$, $h : \mathcal{X} \rightarrow \mathbb{R}^q$, and $g : \mathcal{X} \rightarrow \mathbb{R}^{n \times n}$ are all twice continuously differentiable functions. Define the Lagrangian function $l : \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ by

$$l(x, y, \Psi) = f(x) + \langle y, h(x) \rangle + \langle \Psi, g(x) \rangle, \quad \forall (x, y, \Psi) \in \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}.$$

Then the Karush-Kuhn-Tucher (KKT) condition for the NLNNP is given by

$$\mathcal{J}_x l(x, y, \Psi) = 0, \quad -h(x) = 0, \quad \text{and} \quad \Psi \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(g(x)), \quad (1)$$

where $\mathcal{J}_x l(x, y, \Psi)$ is the derivative of $l(x, y, \Psi)$ at (x, y, Ψ) with respect to $x \in \mathcal{X}$ and $\mathcal{N}_{\mathbb{R}_+^{n \times n}}(a)$ is the normal cone of $\mathbb{R}_+^{n \times n}$ at the point a defined by ([37])

$$\mathcal{N}_{\mathbb{R}_+^{n \times n}}(a) = \begin{cases} \{c \in \mathbb{R}^{n \times n} : \langle c, b - a \rangle \leq 0, \forall b \in \mathbb{R}_+^{n \times n}\} & \text{if } a \in \mathbb{R}_+^{n \times n}, \\ \emptyset & \text{otherwise,} \end{cases}$$

and any point (x, y, Ψ) satisfying (1) is called a KKT point of the NLNNP and the corresponding point x is called a stationary point of the NLNNP.

Suppose that x is a stationary point of the NLNNP. Then there exists a point (y, Ψ) such that (x, y, Ψ) satisfies the KKT condition (1). From [14], it follows that

$$\Psi \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(g(x)) \iff g(x) = \Pi_{\mathbb{R}_+^{n \times n}}(g(x) + \Psi).$$

Hence, the KKT condition (1) can be rewritten as

$$H(x, y, \Psi) := \begin{bmatrix} \mathcal{J}_x l(x, y, \Psi) \\ -h(x) \\ -g(x) + \Pi_{\mathbb{R}_+^{n \times n}}(g(x) + \Psi) \end{bmatrix} = 0, \quad (2)$$

where $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ denotes the metric projection onto $\mathbb{R}_+^{n \times n}$.

We note that we cannot directly use Newton's method to solve (2) since $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ is not continuously differentiable everywhere. Fortunately, one may employ the nonsmooth Newton method for solving (2). To do so, we need the concept of Clarke's generalized Jacobian. We first recall the definitions of Fréchet differentiability and directional differentiability.

Definition 3.1 Let $\Upsilon : \mathcal{X} \rightarrow \mathcal{Y}$ and $x \in \mathcal{X}$.

1). If there exists a linear operator, denoted by $\mathcal{J}_x \Upsilon$, such that

$$\lim_{h \rightarrow 0} \frac{\|\Upsilon(x+h) - \Upsilon(x) - \mathcal{J}_x \Upsilon(h)\|}{\|h\|} = 0$$

for all $h \in \mathcal{X}$, then Υ is Fréchet-differentiable at x and $\mathcal{J}_x \Upsilon$ is the F-derivative of Υ at x .

2). Let $h \in \mathcal{X}$. We define the directional derivative $\Upsilon'(x; h)$ of Υ at x by

$$\Upsilon'(x; h) = \lim_{t \downarrow 0} \frac{\Upsilon(x+th) - \Upsilon(x)}{t}.$$

Υ is said to be directionally differentiable at x if $\Upsilon'(x; h)$ exists for all $h \in \mathcal{X}$.

We now recall the definition of Clarke's generalized Jacobian [11]. Let \mathcal{D} be an open set in \mathcal{Y} and $\Xi : \mathcal{D} \rightarrow \mathcal{Z}$ be a locally Lipschitz continuous function on \mathcal{D} . Using Rademacher's theorem [38, Chap. 9.J], it is easy to know that Ξ is Fréchet differentiable almost everywhere in \mathcal{D} . Then Clarke's generalized Jacobian of Ξ at $y \in \mathcal{D}$ is defined by

$$\partial \Xi(y) := \text{conv}\{\partial_B \Xi(y)\},$$

where “conv” means the convex hull and

$$\partial_B \Xi(y) = \left\{ \lim_{j \rightarrow \infty} \mathcal{J}_{y^j} \Xi(y^j) : y^j \rightarrow y, \Xi \text{ is Fréchet differentiable at } y^j \in \mathcal{D} \right\}.$$

On the generalized Jacobian of composite functions, we have the following result [43, Lemma 2.1].

Lemma 3.2 *Let $\Upsilon : \mathcal{X} \rightarrow \mathcal{Y}$ be a continuously differentiable function on an open neighborhood of $\mathcal{B}(\bar{x})$ of \bar{x} and $\Xi : \mathcal{Y} \rightarrow \mathcal{Z}$ be a locally Lipschitz continuous function on an open set \mathcal{D} containing $\bar{y} := \Upsilon(\bar{x})$. Suppose that Ξ is directionally differentiable at every point in \mathcal{D} and that $\mathcal{J}_x \Upsilon(\bar{x}) : \mathcal{X} \rightarrow \mathcal{Y}$ is onto. Define $\Theta : \mathcal{B}(\bar{x}) \rightarrow \mathcal{Z}$ by $\Theta(x) := \Xi(\Upsilon(x))$, $x \in \mathcal{B}(\bar{x})$. Then we have*

$$\partial_B \Theta(\bar{x}) = \partial_B \Xi(\bar{y}) \mathcal{J}_x \Upsilon(\bar{x}).$$

Next, we review the properties of the metric projection onto a closed convex set. Let $\mathcal{K} \subseteq \mathcal{Z}$ be a closed convex set and $z \in \mathcal{K}$. For any $z \in \mathcal{K}$, let $\Pi_{\mathcal{K}}(z)$ be the metric projection of z onto \mathcal{K} . Then we find the following result on $\partial \Pi_{\mathcal{K}}(z)$ ([27, Proposition 1]).

Lemma 3.3 *Let $\mathcal{K} \subseteq \mathcal{Z}$ be a closed convex set. Then, for any $z \in \mathcal{Z}$ and $V \in \partial \Pi_{\mathcal{K}}(z)$,*

- i) V is self-adjoint.
- ii) $\langle d, Vd \rangle \geq 0$ for all $d \in \mathcal{Z}$.
- iii) $\langle Vd, d - Vd \rangle \geq 0$ for all $d \in \mathcal{Z}$.

For $H : \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n} \rightarrow \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}$ defined in (2), by Lemma 3.2, we can easily derive the following result.

Proposition 3.4 *Let $(\bar{x}, \bar{y}, \bar{\Psi}) \in \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}$. Then, for any $(\Delta x, \Delta y, \Delta \Psi) \in \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}$, we deduce that*

$$\partial H(\bar{x}, \bar{y}, \bar{\Psi})(\Delta x, \Delta y, \Delta \Psi) = \begin{bmatrix} \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x + \mathcal{J}_x h(\bar{x})^* \Delta y + \mathcal{J}_x g(\bar{x})^* \Delta \Psi \\ \mathcal{J}_x h(\bar{x}) \Delta x \\ -\mathcal{J}_x g(\bar{x}) \Delta x + \partial \Pi_{\mathbb{R}_+^{n \times n}}(g(\bar{x}) + \bar{\Psi})(\mathcal{J}_x g(\bar{x}) \Delta x + \Delta \Psi) \end{bmatrix}.$$

From Proposition 3.4, we know that Clarke’s generalized Jacobian of H is given in terms of Clarke’s generalized Jacobian of $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$. In the following, we characterize Clarke’s generalized Jacobian of $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$. Let $\mathbb{R}^{n \times n}$ be equipped with the Frobenius inner product $\langle \cdot, \cdot \rangle$, i.e.,

$$\langle B_1, B_2 \rangle = \text{tr}(B_1^T B_2) \quad \forall B_1, B_2 \in \mathbb{R}^{n \times n},$$

where “tr” denotes the trace of a matrix. Under the Frobenius inner product, the projection $C_+ := \Pi_{\mathbb{R}_+^{n \times n}}(C)$ of a matrix $C \in \mathbb{R}^{n \times n}$ onto $\mathbb{R}_+^{n \times n}$ satisfies the following complementarity condition:

$$\mathbb{R}_+^{n \times n} \ni C_+ \perp (C_+ - C) \in \mathbb{R}_+^{n \times n}, \quad (3)$$

where for any two matrices $C_1, C_2 \in \mathbb{R}^{n \times n}$, $C_1 \perp C_2 \iff \langle C_1, C_2 \rangle = 0$. Define three index sets:

$$\alpha := \{(i, j) \mid C_{ij} > 0\}, \quad \beta := \{(i, j) \mid C_{ij} = 0\}, \quad \gamma := \{(i, j) \mid C_{ij} < 0\}.$$

Then

$$C = P(C_\alpha) + P(C_\beta) + P(C_\gamma). \quad (4)$$

Define the matrix $U(C) \in \mathbb{R}^{n \times n}$ by

$$U_{ij}(C) := \frac{\max\{C_{ij}, 0\}}{|C_{ij}|}, \quad i, j = 1, \dots, n, \quad (5)$$

where $0/0$ is defined to be 1. It is easy to check that, for any $H \in \mathbb{R}^{n \times n}$, the directional derivative $\Pi'_{\mathbb{R}_+^{n \times n}}(C; H)$ is given by

$$\Pi'_{\mathbb{R}_+^{n \times n}}(C; H) = P(U_{\alpha \cup \gamma}(C) \circ H_{\alpha \cup \gamma}) + P(\Pi_{\mathbb{R}_+^{|\beta|}}(H_\beta)), \quad (6)$$

where \circ denotes the Hadamard product.

Based on the previous analysis, we can easily derive the following differential properties of $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$.

Proposition 3.5 *Let the absolute value function $|\cdot|_{\mathbb{R}^{m \times n}}$ be defined by*

$$|C|_{\mathbb{R}^{m \times n}} = \Pi_{\mathbb{R}_+^{m \times n}}(C) + \Pi_{\mathbb{R}_+^{m \times n}}(-C) \quad \forall C \in \mathbb{R}^{m \times n}.$$

- (a) $|\cdot|_{\mathbb{R}^{n \times n}}$ and $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ are F -differentiable at $C \in \mathbb{R}^{n \times n}$ if and only if C has no zero entry. In this case, $\Pi'_{\mathbb{R}_+^{n \times n}}(C; \cdot) = U(C) \circ$, where $U(C)$ is defined in (5).
- (b) For any $C \in \mathbb{R}^{n \times n}$, $\Pi'_{\mathbb{R}_+^{n \times n}}(C, \cdot)$ is F -differentiable at $H \in \mathbb{R}^{n \times n}$ if and only if H_β has no zero entry.
- (c) For any $C, H \in \mathbb{R}^{n \times n}$, $\Pi'_{\mathbb{R}_+^{n \times n}}(C; H)$ is given by (6).

To characterize Clarke's generalized Jacobian of $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$, we need the following lemma.

Lemma 3.6 *For any $C \in \mathbb{R}^{n \times n}$, define $\Theta(\cdot) := \Pi'_{\mathbb{R}_+^{n \times n}}(C; \cdot)$. Then*

$$\partial_B \Pi_{\mathbb{R}_+^{n \times n}}(C) = \partial_B \Theta(0).$$

Proof: Let $V \in \partial_B \Pi_{\mathbb{R}_+^{n \times n}}(C)$. By Proposition 3.5 and the definition of the elements in $\partial_B \Pi_{\mathbb{R}_+^{n \times n}}(C)$, it follows that there exists a sequence of no-zero-entry matrices $\{C^k\}$ in $\mathbb{R}^{n \times n}$ converging to C such that $V = \lim_{k \rightarrow \infty} \mathcal{J}_{C^k} \Pi_{\mathbb{R}_+^{n \times n}}(C^k)$. Notice that $C = \lim_{k \rightarrow \infty} C^k$ implies that for all k large enough, we have $C_{\alpha \cup \gamma}^k$ has no zero entry and $\lim_{k \rightarrow \infty} C_\beta^k = 0$. For any $H \in \mathbb{R}^{n \times n}$, we have

$$Q^k := \mathcal{J}_{C^k} \Pi_{\mathbb{R}_+^{n \times n}}(C^k)(H) = U(C^k) \circ H,$$

which implies $|C^k| \circ Q^k = C_+^k \circ H$, i.e.,

$$P(|C_{\alpha\cup\gamma}^k| \circ Q_{\alpha\cup\gamma}^k) + P(|C_{\beta}^k| \circ Q_{\beta}^k) = P((C_{\alpha\cup\gamma}^k)_+ \circ H_{\alpha\cup\gamma}) + P((C_{\beta}^k)_+ \circ H_{\beta}).$$

Thus, $Q_{\alpha\cup\gamma}^k = U_{\alpha\cup\gamma}(C^k) \circ H_{\alpha\cup\gamma}$ and $Q_{\beta}^k = U_{\beta}(C^k) \circ H_{\beta}$. By taking a subsequence if necessary, we may assume that $\{Q_{\beta}^k\}$ is a convergent sequence. Then, for any $H \in \mathbb{R}^{n \times n}$, we deduce that

$$V(H) = P(U_{\alpha\cup\gamma}(C) \circ H_{\alpha\cup\gamma}) + P\left(\lim_{k \rightarrow \infty} \{U_{\beta}(C^k) \circ H_{\beta}\}\right). \quad (7)$$

Let $R^k := P(C_{\beta}^k)$. Since C_{β}^k has no zero entry, we know that Θ is F-differentiable at R^k and for any $H \in \mathbb{R}^{n \times n}$, we have

$$\begin{aligned} \mathcal{J}_{R^k}\Theta(R^k)(H) &= \lim_{t \downarrow 0} \frac{\Theta(P(C_{\beta}^k) + tH) - \Theta(P(C_{\beta}^k))}{t} \\ &= P(U_{\alpha\cup\gamma}(C) \circ H_{\alpha\cup\gamma}) + P\left(\lim_{t \downarrow 0} \frac{\Pi_{\mathbb{R}_+^{|\beta|}}(C_{\beta}^k + tH_{\beta}) - \Pi_{\mathbb{R}_+^{|\beta|}}(C_{\beta}^k)}{t}\right) \\ &= P(U_{\alpha\cup\gamma}(C) \circ H_{\alpha\cup\gamma}) + P(U_{\beta}(C^k) \circ H_{\beta}), \end{aligned}$$

where the second equality uses (6) and the third equality uses Proposition 3.5 (a) applied to $\Pi_{\mathbb{R}_+^{|\beta|}}$ at the F-differentiable point C_{β}^k . The latter, together with (7), implies that $V(H) = \lim_{k \rightarrow \infty} \mathcal{J}_{R^k}\Theta(R^k)(H)$. By the arbitrariness of $H \in \mathbb{R}^{n \times n}$, we have $V \in \partial_B\Theta(0)$.

Conversely, let $V \in \partial_B\Theta(0)$. Notice that Θ is F-differentiable at $R \in \mathbb{R}^{n \times n}$ if and only if R_{β} has no zero entry. Then there exists a sequence of matrices $\{R^k\}$ in $\mathbb{R}^{n \times n}$ converges to 0 such that R_{β}^k has no zero entry for every k and $V = \lim_{k \rightarrow \infty} \mathcal{J}_{R^k}\Theta(R^k)$. For any $H \in \mathbb{R}^{n \times n}$, we infer that

$$\begin{aligned} \mathcal{J}_{R^k}\Theta(R^k)(H) &= \lim_{t \downarrow 0} \frac{\Theta(R^k + tH) - \Theta(R^k)}{t} \\ &= P(U_{\alpha\cup\gamma}(C) \circ H_{\alpha\cup\gamma}) + P\left(\lim_{t \downarrow 0} \frac{\Pi_{\mathbb{R}_+^{|\beta|}}(R_{\beta}^k + tH_{\beta}) - \Pi_{\mathbb{R}_+^{|\beta|}}(R_{\beta}^k)}{t}\right) \\ &= P(U_{\alpha\cup\gamma}(C) \circ H_{\alpha\cup\gamma}) + P(U_{\beta}(R^k) \circ H_{\beta}). \end{aligned}$$

Let $C^k := C + P(R_{\beta}^k) = P(C_{\alpha\cup\gamma}) + P(R_{\beta}^k)$. It is obvious that C^k has no zero entry for every k and

$$|C^k|_{\mathbb{R}^{n \times n}} = P(|C_{\alpha\cup\gamma}|) + P(|R_{\beta}^k|) \quad \text{and} \quad (C^k)_+ = P((C_{\alpha\cup\gamma})_+) + P(\Pi_{\mathbb{R}_+^{|\beta|}}R_{\beta}^k).$$

Hence, $\Pi_{\mathbb{R}_+^{n \times n}}$ is differentiable at C^k and for any $H \in \mathbb{R}^{n \times n}$,

$$Q^k := \mathcal{J}_{C^k}\Pi_{\mathbb{R}_+^{n \times n}}(C^k)(H) = U(C^k) \circ H,$$

which leads to $|C^k|_{\mathbb{R}^{n \times n}} \circ Q^k = (C^k)_+ \circ H$. Thus,

$$P(|C_{\alpha\cup\gamma}| \circ Q_{\alpha\cup\gamma}^k) + P(|R_{\beta}^k| \circ Q_{\beta}^k) = P((C_{\alpha\cup\gamma})_+ \circ H_{\alpha\cup\gamma}) + P((R_{\beta}^k)_+ \circ H_{\beta}),$$

which gives rise to

$$Q_{\alpha \cup \gamma}^k = U_{\alpha \cup \gamma}(C) \circ H_{\alpha \cup \gamma} \quad \text{and} \quad Q_{\beta}^k = U_{\beta}(P(C^k)) \circ H_{\beta}.$$

Therefore, $V(H) = \lim_{k \rightarrow \infty} \mathcal{J}_{C^k} \Pi_{\mathbb{R}_+^{n \times n}}(C^k)(H)$ and then $V \in \partial_B \Pi_{\mathbb{R}_+^{n \times n}}(C)$. The proof is complete. \square

We are now ready for establishing the following result on $\partial \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$.

Proposition 3.7 *Suppose that $C \in \mathbb{R}^{n \times n}$ has the decomposition as in (4). Then for any $V \in \partial_B \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ ($\partial \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$, respectively), there exists a $W \in \partial_B \Pi_{\mathbb{R}_+^{|\beta|}}(0)$ ($\partial \Pi_{\mathbb{R}_+^{|\beta|}}(0)$, respectively) such that*

$$V(H) = P(U_{\alpha \cup \gamma}(C) \circ H_{\alpha \cup \gamma}) + P(W(H_{\beta})). \quad (8)$$

Conversely, for any $W \in \partial_B \Pi_{\mathbb{R}_+^{|\beta|}}(0)$ ($\partial \Pi_{\mathbb{R}_+^{|\beta|}}(0)$, respectively), there exists a $V \in \partial_B \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ ($\partial \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$, respectively) such that (8) holds.

Proof: We only need to prove (8) for $V \in \partial_B \Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ and $W \in \partial_B \Pi_{\mathbb{R}_+^{|\beta|}}(0)$.

Let $\Theta(\cdot) := \Pi'_{\mathbb{R}_+^{n \times n}}(C; \cdot)$. By (6), we have

$$\Theta(H) = P(U_{\alpha \cup \gamma}(C) \circ H_{\alpha \cup \gamma}) + P(\Pi_{\mathbb{R}_+^{|\beta|}}(H_{\beta})), \quad \forall H \in \mathbb{R}^{n \times n}.$$

By Lemma 3.6, we obtain (8). \square

We now use Clarke's generalized Jacobian-based Newton method for solving the nonsmooth equation (2):

$$(x^{j+1}, y^{j+1}, \Psi^{j+1}) = (x^j, y^j, \Psi^j) - V_j^{-1} H(x^j, y^j, \Psi^j), \quad V_j \in \partial H(x^j, y^j, \Psi^j), \quad j = 0, 1, 2, \dots \quad (9)$$

To guarantee the superlinear convergence of (9), we need the semismoothness of H , whose notion is formally defined below [28, 34].

Definition 3.8 *Let $\Upsilon : \mathcal{D} \subseteq \mathcal{Y} \rightarrow \mathcal{Z}$ be a locally Lipschitz continuous function on the open set \mathcal{D} .*

- 1) Υ is said to be *semismooth* at $y \in \mathcal{D}$ if Υ is directionally differentiable at y and for any $V \in \partial \Upsilon(y + h)$ and $h \rightarrow 0$,

$$\Upsilon(y + h) - \Upsilon(y) - V(h) = o(\|h\|).$$

- 2) Υ is said to be *strongly semismooth* at $y \in \mathcal{D}$ if Υ is semismooth at y and for any $V \in \partial \Upsilon(y + h)$ and $h \rightarrow 0$,

$$\Upsilon(y + h) - \Upsilon(y) - V(h) = O(\|h\|^2).$$

Regarding the superlinear convergence of (9), we state the following result [34, Theorem 3.2].

Proposition 3.9 *Let $H : \mathcal{O} \subseteq \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n} \rightarrow \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}$ be a locally Lipschitz continuous function on the open set \mathcal{O} . Let $(\bar{x}, \bar{y}, \bar{\Psi}) \in \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}$ be a solution to (2). Suppose that H is semismooth at $(\bar{x}, \bar{y}, \bar{\Psi})$ and any element in $\partial H(\bar{x}, \bar{y}, \bar{\Psi})$ is nonsingular. Then every sequence generated by (9) converges to $(\bar{x}, \bar{y}, \bar{\Psi})$ superlinearly provided the initial point (x^0, y^0, Ψ^0) is sufficiently close to $(\bar{x}, \bar{y}, \bar{\Psi})$. Moreover, the convergence rate is quadratic if H is strongly semismooth at $(\bar{x}, \bar{y}, \bar{\Psi})$.*

In order to solve (2) by using the nonsmooth Newton method (9), the two assumptions in Proposition 3.9 should be satisfied. The strong semismoothness of H holds since the metric projection $\Pi_{\mathbb{R}_+^{n \times n}}(\cdot)$ is strongly semismooth. In what follows, we explore the nonsingularity conditions of Clarke's generalized Jacobian $\partial H(\cdot)$. We need the concepts of strong second order sufficient condition and constraint nondegeneracy for the NLNNP. Let $\mathcal{K} \subseteq \mathcal{Z}$ be a closed convex set and $z \in \mathcal{K}$. Define

$$\text{dist}(z, \mathcal{K}) := \inf\{\|z - d\| : d \in \mathcal{K}\}.$$

Then the *tangent cone* $\mathcal{T}_{\mathcal{K}}(z)$ of \mathcal{K} at $z \in \mathcal{K}$ is given by [3, §2.2.4])

$$\mathcal{T}_{\mathcal{K}}(z) := \{d \in \mathcal{K} : \text{dist}(z + td, \mathcal{K}) = o(t), t \geq 0\}.$$

For any $z \in \mathcal{K}$, let $\text{lin}(\mathcal{T}_{\mathcal{K}}(z))$ denote the linearity space of $\mathcal{T}_{\mathcal{K}}(z)$, i.e., the largest linear space in $\mathcal{T}_{\mathcal{K}}(z)$. By (6), we have

$$\mathcal{T}_{\mathbb{R}_+^{n \times n}}(C_+) = \{H \in \mathbb{R}^{n \times n} \mid H = \Pi'_{\mathbb{R}_+^{n \times n}}(C_+; H)\} = \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\} \quad (10)$$

and

$$\text{lin}(\mathcal{T}_{\mathbb{R}_+^{n \times n}}(C_+)) = \{H \in \mathbb{R}^{n \times n} \mid H_\beta = 0, H_\gamma = 0\}. \quad (11)$$

The following definition is the constraint nondegeneracy for the NLNNP, which is originally introduced by Robinson [36].

Definition 3.10 *We say that a feasible point \bar{x} to the NLNNP is constraint nondegenerate if*

$$\begin{pmatrix} \mathcal{J}_x h(\bar{x}) \\ \mathcal{J}_x g(\bar{x}) \end{pmatrix} \mathcal{X} + \begin{pmatrix} \{0\} \\ \text{lin}(\mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x}))) \end{pmatrix} = \begin{pmatrix} \mathbb{R}^q \\ \mathbb{R}^{n \times n} \end{pmatrix}. \quad (12)$$

As noted in [36] and [41], the constraint nondegeneracy for the NLNNP is equivalent to the stronger linear independence constraint qualification (LICQ) (cf., [30]):

$$\{\mathcal{J}_x h_j(\bar{x})\}_{j=1}^q \text{ and } \{\mathcal{J}_x g_{ij}(\bar{x})\}_{(i,j) \in \mathcal{A}(\bar{x})} \text{ is linearly independent.}$$

Moreover, in this case, the set $\mathcal{M}(\bar{x})$ of Lagrangian multipliers defined by

$$\mathcal{M}(\bar{x}) := \{(\bar{y}, \bar{\Psi}) \mid (\bar{x}, \bar{y}, \bar{\Psi}) \text{ is a KKT point of the NLNNP}\}$$

is a singleton.

In the following, we give the concept of strong second order sufficient condition for the NLNNP. We first need to characterize the critical cone of the NLNNP. The critical cone of $\mathbb{R}_+^{n \times n}$ at $C \in \mathbb{R}^{n \times n}$ associated with the complementarity problem (3) is defined by

$$\mathcal{C}(C; \mathbb{R}_+^{n \times n}) := \mathcal{T}_{\mathbb{R}_+^{n \times n}}(C_+) \cap (C_+ - C)^\perp,$$

i.e.,

$$\mathcal{C}(C; \mathbb{R}_+^{n \times n}) = \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\}. \quad (13)$$

Hence, the affine hull $\text{aff}(\mathcal{C}(C; \mathbb{R}_+^{n \times n}))$ of $\mathcal{C}(C; \mathbb{R}_+^{n \times n})$ is given by

$$\text{aff}(\mathcal{C}(C; \mathbb{R}_+^{n \times n})) = \{H \in \mathbb{R}^{n \times n} \mid H_\gamma = 0\}. \quad (14)$$

Assume that \bar{x} be a stationary point of the NLNNP. Then $\mathcal{M}(\bar{x})$ is nonempty. The critical cone $\mathcal{C}(\bar{x})$ of the NLNNP at \bar{x} is defined by

$$\begin{aligned} \mathcal{C}(\bar{x}) &:= \{d \mid \mathcal{J}_x h(\bar{x})d = 0, \mathcal{J}_x g(\bar{x})d \in \mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})), \mathcal{J}_x f(\bar{x})d \leq 0\} \\ &= \{d \mid \mathcal{J}_x h(\bar{x})d = 0, \mathcal{J}_x g(\bar{x})d \in \mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})), \mathcal{J}_x f(\bar{x})d = 0\}. \end{aligned}$$

Since \bar{x} is a stationary point of the NLNNP, there exists $(\bar{y}, \bar{\Psi}) \in \mathcal{M}(\bar{x})$ such that

$$\mathcal{J}_x l(\bar{x}, \bar{y}, \bar{\Psi}) = 0, \quad -h(\bar{x}) = 0, \quad \text{and} \quad \bar{\Psi} \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})).$$

From [14], we have

$$\bar{\Psi} \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})) \iff \mathbb{R}_+^{n \times n} \ni (-\bar{\Psi}) \perp g(\bar{x}) \in \mathbb{R}_+^{n \times n}.$$

Thus $C := g(\bar{x}) + \bar{\Psi}$ takes the form reported in (4):

$$g(\bar{x}) = P(C_{\alpha \cup \beta}) \quad \text{and} \quad \bar{\Psi} = P(C_{\beta \cup \gamma}). \quad (15)$$

By (10) and (11), we get

$$\begin{aligned} \mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})) &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\}, \\ \mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})) \cap \bar{\Psi}^\perp &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\}, \\ \text{lin}(\mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x}))) &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta = 0, H_\gamma = 0\}. \end{aligned} \quad (16)$$

Since $\mathcal{M}(g(\bar{x}))$ is nonempty,

$$\begin{aligned} \mathcal{C}(\bar{x}) &= \{d \mid \mathcal{J}_x h(\bar{x})d = 0, (\mathcal{J}_x g(\bar{x})d)_\beta \geq 0, (\mathcal{J}_x g(\bar{x})d)_\gamma = 0\} \\ &= \{d \mid \mathcal{J}_x h(\bar{x})d = 0, \mathcal{J}_x g(\bar{x})d \in \mathcal{C}(C; \mathbb{R}_+^{n \times n})\}, \end{aligned}$$

where $\mathcal{C}(C; \mathbb{R}_+^{n \times n})$ is the critical cone of $\mathbb{R}_+^{n \times n}$ at $C := g(\bar{x}) + \bar{\Psi}$. It is difficult to determine the affine hull $\text{aff}(\mathcal{C}(\bar{x}))$ of $\mathcal{C}(\bar{x})$. However, based on $(\bar{y}, \bar{\Psi})$, we can provide an approximation by

$$\text{app}(\bar{y}, \bar{\Psi}) := \{d \mid \mathcal{J}_x h(\bar{x})d = 0, \mathcal{J}_x g(\bar{x})d \in \text{aff}(\mathcal{C}(C; \mathbb{R}_+^{n \times n}))\}.$$

which, together with (14), leads to

$$\text{app}(\bar{y}, \bar{\Psi}) := \{d \mid \mathcal{J}_x h(\bar{x})d = 0, (\mathcal{J}_x g(\bar{x})d)_\gamma = 0\}. \quad (17)$$

On the relations between $\text{app}(\bar{y}, \bar{\Psi})$ and $\text{aff}(\mathcal{C}(\bar{x}))$, we can easily deduce the following result.

Proposition 3.11 *Suppose that there exists a direction $d \in \mathcal{C}(\bar{x})$ such that $(\mathcal{J}_x g(\bar{x})d)_\beta > 0$. Then we have*

$$\text{aff}(\mathcal{C}(\bar{x})) = \text{app}(\bar{y}, \bar{\Psi}).$$

The following definition is a strong second order sufficient condition for the NLNNP, which is slightly different from the original version introduced by Robinson [35].

Definition 3.12 *Let $(\bar{x}, \bar{y}, \bar{\Psi})$ be a KKT point of the NLNNP. We say that the strong second order sufficient condition holds at \bar{x} if*

$$\langle d, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi})d \rangle > 0 \quad \forall d \in \text{aff}(\bar{y}, \bar{\Psi}) \setminus \{0\}. \quad (18)$$

To show the nonsingularity of Clarke's generalized Jacobian of H , we need the following useful result.

Proposition 3.13 *Suppose that $G \in \mathbb{R}_+^{n \times n}$ and $\Psi \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(Y)$. Then for any $V \in \partial \Pi_{\mathbb{R}_+^{n \times n}}(G + \Psi)$ and $\Delta G, \Delta \Psi \in \mathbb{R}^{n \times n}$ such that $\Delta G = V(\Delta G + \Delta \Psi)$, it holds that*

$$\langle \Delta G, \Delta \Psi \rangle \geq 0.$$

Proof: Let $C := G + \Psi$. We have from [14] that $G = \Pi_{\mathbb{R}_+^{n \times n}}(G + \Psi) = \Pi_{\mathbb{R}_+^{n \times n}}(C)$ and $\langle G, \Psi \rangle = 0$. Hence,

$$G = P(C_{\alpha \cup \beta}) \quad \text{and} \quad \Psi = P(C_{\beta \cup \gamma}).$$

By Proposition 3.7, there exists a $W \in \partial \Pi_{|\beta|}(0)$ such that

$$V(\Delta G + \Delta \Psi) = P(U_{\alpha \cup \gamma}(C) \circ (\Delta G_{\alpha \cup \gamma} + \Delta \Psi_{\alpha \cup \gamma})) + P(W(\Delta G_\beta + \Delta \Psi_\beta)).$$

Taking into consideration the assumption that $\Delta G = V(\Delta G + \Delta \Psi)$, we infer that

$$\Delta \Phi_\alpha = 0, \quad \Delta G_\gamma = 0, \quad \Delta G_\beta = W(\Delta G_\beta + \Delta \Psi_\beta). \quad (19)$$

From Lemma 3.3 iii) and (19), we obtain

$$\begin{aligned} \langle \Delta G, \Delta \Psi \rangle &= \langle \Delta G_\alpha, \Delta \Psi_\alpha \rangle + \langle \Delta G_\beta, \Delta \Psi_\beta \rangle + \langle \Delta G_\gamma, \Delta \Psi_\gamma \rangle \\ &= \langle \Delta G_\beta, \Delta \Psi_\beta \rangle = \langle W(\Delta G_\beta + \Delta \Psi_\beta), \Delta \Psi_\beta \rangle \\ &= \langle W(\Delta G_\beta + \Delta \Psi_\beta), (\Delta G_\beta + \Delta \Psi_\beta) - W(\Delta G_\beta + \Delta \Psi_\beta) \rangle \geq 0. \end{aligned}$$

□

We are now ready to state our result on the nonsingularity of Clarke's generalized Jacobian of the mapping H defined in (2).

Theorem 3.14 *Let \bar{x} be a feasible solution to the NLNNP. Let $(\bar{y}, \bar{\Psi})$ satisfies $H(\bar{x}, \bar{y}, \bar{\Psi}) = 0$. Suppose that the strong second order sufficient condition (18) holds at \bar{x} and \bar{x} is constraint nondegenerate. Then every element in $\partial H(\bar{x}, \bar{y}, \bar{\Psi})$ is nonsingular.*

Proof: Let V be an arbitrary element in $\partial H(\bar{x}, \bar{y}, \bar{\Psi})$. To show that V is nonsingular, we only need to prove that

$$V(\Delta x, \Delta y, \Delta \Psi) = 0 \quad \forall (\Delta x, \Delta y, \Delta \Psi) \in \mathcal{X} \times \mathbb{R}^q \times \mathbb{R}^{n \times n}.$$

Let $C := g(\bar{x}) + \bar{\Psi}$. Then C , $g(\bar{x})$ and $\bar{\Psi}$ take the forms of (4) and (15), respectively. By Lemma 3.2, there exists an element $W \in \partial \Pi_{\mathbb{R}_+^{n \times n}}(C)$ such that

$$V(\Delta x, \Delta y, \Delta \Psi) = \begin{bmatrix} \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x + \mathcal{J}_x h(\bar{x})^* \Delta y + \mathcal{J}_x g(\bar{x})^* \Delta \Psi \\ \mathcal{J}_x h(\bar{x}) \Delta x \\ -\mathcal{J}_x g(\bar{x}) \Delta x + W(\mathcal{J}_x g(\bar{x}) \Delta x + \Delta \Psi) \end{bmatrix} = 0. \quad (20)$$

By Proposition 3.7, (17), and the second and the third equations of (20), we obtain

$$\Delta x \in \text{app}(\bar{y}, \bar{\Psi}). \quad (21)$$

It follows from the first and second equations of (20) that

$$\begin{aligned} 0 &= \langle \Delta x, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x + \mathcal{J}_x h(\bar{x})^* \Delta y + \mathcal{J}_x g(\bar{x})^* \Delta \Psi \rangle \\ &= \langle \Delta x, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x \rangle + \langle \Delta x, \mathcal{J}_x h(\bar{x})^* \Delta y \rangle + \langle \Delta x, \mathcal{J}_x g(\bar{x})^* \Delta \Psi \rangle \\ &= \langle \Delta x, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x \rangle + \langle \Delta y, \mathcal{J}_x h(\bar{x}) \Delta x \rangle + \langle \Delta \Psi, \mathcal{J}_x g(\bar{x}) \Delta x \rangle \\ &= \langle \Delta x, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x \rangle + \langle \Delta \Psi, \mathcal{J}_x g(\bar{x}) \Delta x \rangle. \end{aligned} \quad (22)$$

By Proposition 3.13 and the third equation of (20), we find

$$\langle \Delta \Psi, \mathcal{J}_x g(\bar{x}) \Delta x \rangle \geq 0.$$

The latter, together with (22), implies that

$$0 \geq \langle \Delta x, \mathcal{J}_{xx}^2 l(\bar{x}, \bar{y}, \bar{\Psi}) \Delta x \rangle. \quad (23)$$

It follows from (23), (21) and the strong second order sufficient condition (18) that

$$\Delta x = 0. \quad (24)$$

Then (20) reduces to

$$\begin{bmatrix} \mathcal{J}_x h(\bar{x})^* \Delta y + \mathcal{J}_x g(\bar{x})^* \Delta \Psi \\ W(\Delta \Psi) \end{bmatrix} = 0. \quad (25)$$

By Proposition 3.7 and the second equation of (25), we obtain

$$(\Delta \Psi)_\alpha = 0. \quad (26)$$

By the constraint nondegeneracy condition (12), there exist $d \in \mathcal{X}$ and $Q \in \text{lin}(\mathcal{T}_{\mathbb{R}_+^{n \times n}}(g(\bar{x})))$ such that

$$\mathcal{J}_x h(\bar{x}) d = \Delta y \quad \text{and} \quad \mathcal{J}_x g(\bar{x}) d + Q = \Delta \Psi,$$

which, together with the first equation of (25), implies that

$$\begin{aligned}
\langle \Delta y, \Delta y \rangle + \langle \Delta \Psi, \Delta \Psi \rangle &= \langle \mathcal{J}_x h(\bar{x})d, \Delta y \rangle + \langle \mathcal{J}_x g(\bar{x})d + Q, \Delta \Psi \rangle \\
&= \langle d, \mathcal{J}_x h(\bar{x})^* \Delta y \rangle + \langle d, \mathcal{J}_x g(\bar{x})^* \Delta \Psi \rangle + \langle Q, \Delta \Psi \rangle \\
&= \langle d, \mathcal{J}_x h(\bar{x})^* \Delta y + \mathcal{J}_x g(\bar{x})^* \Delta \Psi \rangle + \langle Q, \Delta \Psi \rangle \\
&= \langle Q, \Delta \Psi \rangle.
\end{aligned} \tag{27}$$

From (16), (26), and (27), we get

$$\langle \Delta y, \Delta y \rangle + \langle \Delta \Psi, \Delta \Psi \rangle = 0.$$

Hence,

$$\Delta y = 0 \quad \text{and} \quad \Delta \Psi = 0. \tag{28}$$

We see from (24) and (28) that V is nonsingular. \square

Finally, we define the monotonicity and related concepts for matrix-valued functions, which were originally introduced for vector-valued functions [4, 16, 48].

Definition 3.15 *Let \mathcal{X}^n denote the space of real $n \times n$ matrices or the space of real symmetric $n \times n$ matrices, which is equipped with the Frobenius inner product $\langle \cdot, \cdot \rangle$ and the related induced Frobenius norm $\| \cdot \|$.*

1). $\Upsilon : \mathcal{X}^n \rightarrow \mathcal{X}^n$ is a monotone function if

$$\langle Y - Z, \Upsilon(Y) - \Upsilon(Z) \rangle \geq 0 \quad \text{for all } Y, Z \in \mathcal{X}^n.$$

2). $\Upsilon : \mathcal{X}^n \rightarrow \mathcal{X}^n$ is a P_0 -function if there exists an index (i, j) such that

$$Y_{ij} \neq Z_{ij} \quad \text{and} \quad (Y_{ij} - Z_{ij})(\Upsilon_{ij}(Y) - \Upsilon_{ij}(Z)) \geq 0 \quad \text{for all } Y, Z \in \mathcal{X}^n \text{ and } Y \neq Z.$$

3). $\Upsilon : \mathcal{X}^n \rightarrow \mathcal{X}^n$ is a P -function if there exists an index (i, j) such that

$$(Y_{ij} - Z_{ij})(\Upsilon_{ij}(Y) - \Upsilon_{ij}(Z)) > 0 \quad \text{for all } Y, Z \in \mathcal{X}^n \text{ and } Y \neq Z.$$

We observe that every monotone matrix-valued function is also a P_0 -function.

3.2 A Nonsmooth Newton-Type Method

In this section, we present a globalized nonsmooth Newton-type method for solving the NIEP. Given the eigendata $(\Lambda, X) \in \mathbb{R}^{n \times p} \times \mathbb{R}^{p \times p}$ as in Section 2, the NIEP is to find a matrix $A \in \mathbb{R}_+^{n \times n}$ such that

$$AX = X\Lambda. \tag{29}$$

Let $K = X^T$ and $B = (X\Lambda)^T$. We note that $\bar{A} \in \mathbb{R}^{n \times n}$ is a solution to (29) if and only if $\bar{Y} = \bar{A}^T \in \mathbb{R}^{n \times n}$ is a global solution of the following convex optimization problem

$$\begin{cases} \min & f(Y) := \frac{1}{2} \|KY - B\|^2 \\ \text{s.t.} & Y \in \mathbb{R}_+^{n \times n} \end{cases} \quad (30)$$

with $f(\bar{Y}) = 0$. The well-known KKT condition for Problem (30) is given by

$$\mathcal{J}_Y f(Y) - Z = 0, \quad Y \in \mathbb{R}_+^{n \times n}, \quad Z \in \mathbb{R}_+^{n \times n}, \quad \langle Y, Z \rangle = 0, \quad (31)$$

From [14], we have

$$Y \in \mathbb{R}_+^{n \times n}, Z \in \mathbb{R}_+^{n \times n}, \langle Y, Z \rangle = 0 \iff -Z \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(Y) \iff Y = \Pi_{\mathbb{R}_+^{n \times n}}(Y - Z).$$

Thus, the KKT condition (31) can be written as

$$H(Y, Z) := \begin{bmatrix} \mathcal{J}_Y f(Y) - Z \\ -Y + \Pi_{\mathbb{R}_+^{n \times n}}(Y - Z) \end{bmatrix} = 0. \quad (32)$$

For Problem (30), it is obvious that the LICQ is satisfied at \bar{Y} . Therefore,

$$\mathcal{M}(\bar{Y}) = \{\bar{Z} \mid (\bar{Y}, \bar{Z}) \text{ is a KKT point of Problem (30)}\}$$

is singleton.

As in Subsection 3.1, one may use Clarke's generalized Jacobian-based Newton method for solving (32), where the unknown variables Y and Z are to be determined. Sparked by [13, 48], in this paper, we solve Problem (30) by constructing an equivalent nonsmooth equation to the KKT condition (31). Let

$$F(Y) := \mathcal{J}_Y f(Y) = K^T(KY - B). \quad (33)$$

Then the KKT condition (31) is reduced to the following MCP

$$Y \in \mathbb{R}_+^{n \times n}, \quad F(Y) \in \mathbb{R}_+^{n \times n}, \quad \langle Y, F(Y) \rangle = 0. \quad (34)$$

By using the well-known Fischer function [17, 19] defined by

$$\omega(a, b) := \sqrt{a^2 + b^2} - (a + b) \quad \forall a, b \in \mathbb{R}, \quad (35)$$

which has an important property:

$$\omega(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0,$$

Solving the MCP (34) is equivalent to the solution of the following nonsmooth equation

$$\Phi(Y) = 0, \quad (36)$$

where $\Phi_{ij}(Y) := \omega(Y_{ij}, F_{ij}(Y))$ for $i, j = 1, \dots, n$. Also, define the merit function $\phi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ by

$$\phi(Y) := \frac{1}{2} \|\Phi(Y)\|^2. \quad (37)$$

In what follows, we propose a Newton-type method for solving the nonsmooth equation (36). We first show the monotonicity of the matrix-valued function F in (33).

Proposition 3.16 *The function $F : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ defined in (33) is monotone matrix-valued function.*

Proof: By (33), for any $Y_1, Y_2 \in \mathbb{R}^{n \times n}$, we find

$$\langle Y_1 - Y_2, F(Y_1) - F(Y_2) \rangle = \langle Y_1 - Y_2, K^T K(Y_1 - Y_2) \rangle = \langle K(Y_1 - Y_2), K(Y_1 - Y_2) \rangle \geq 0.$$

This shows that F is monotone. □

We note that F defined in (33) is continuously differentiable. Then we have the following result on Clarke's generalized Jacobian of Φ defined in (36).

Proposition 3.17 *Let $Y \in \mathbb{R}^{n \times n}$ and Φ be defined in (36). Then, for any $H \in \mathbb{R}^{n \times n}$,*

$$\partial\Phi(Y)H \subseteq (\Gamma(Y) - E) \circ H + (\Omega(Y) - E) \circ (\mathcal{J}_Y F(H)),$$

where E is the n -by- n matrix of all ones and $\Gamma(Y)$ and $\Omega(Y)$ are two n -by- n matrices with entries determined by

$$\Gamma_{ij}(Y) = \frac{Y_{ij}}{\|(Y_{ij}, F_{ij}(Y))\|}, \quad \Omega_{ij}(Y) = \frac{F_{ij}(Y)}{\|(Y_{ij}, F_{ij}(Y))\|},$$

if $(Y_{ij}, F_{ij}(Y)) \neq (0, 0)$ and by

$$\Gamma_{ij}(Y) = \Gamma_{ij}, \quad \Omega_{ij}(Y) = \Omega_{ij} \quad \text{for every } (\Gamma_{ij}, \Omega_{ij}) \text{ such that } \|(\Gamma_{ij}, \Omega_{ij})\| \leq 1$$

if $Y_{ij} = 0 = F_{ij}(Y)$.

Proof: It follows from [11, page 75] and the differentiability of F . □

One important property of the function Φ defined in (36) is its strong semismoothness.

Proposition 3.18 *The function Φ is strongly semismooth.*

Proof: The function Φ is strongly semismooth if and only if its each component is strongly semismooth [44]. Notice that $\Phi_{ij}(Y)$ is a composite of the function $\omega : \mathbb{R}^2 \rightarrow \mathbb{R}$ and the continuously differentiable function $(Y_{ij}, F_{ij}(Y)) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^2$. Since the derivative of F at Y is a constant, we can show the strong semismoothness of $\Phi_{ij}(Y)$ as in [33, Lemma 3.1]. □

The following proposition furnishes two properties of the function ϕ in (37).

Proposition 3.19 a). *The function ϕ defined in (37) is continuously differentiable and its gradient at $Y \in \mathbb{R}^{n \times n}$ is given by*

$$\nabla\phi(Y) = \{V^* \Phi(Y) \mid V \in \partial\Phi(Y)\} \subseteq (\Gamma(Y) - E) \circ \Phi(Y) + (\mathcal{J}_Y F)^*((\Omega(Y) - E) \circ \Phi(Y)).$$

b). *Any stationary point of ϕ is a solution to Problem (30).*

Proof: We first prove part a). By using the generalized Jacobian chain rule in [11, Theorem 2.6.6], we get $\partial\phi(Y) = \{V^*\Phi(Y) \mid V \in \partial\Phi(Y)\}$. If $Y_{ij} = 0 = F_{ij}(Y)$, then $\Phi_{ij}(Y) = 0$. In this case, the multivalued entries of $\partial\phi(Y)$ are canceled by the zero components of $\Phi(Y)$. Then, $\partial\phi(Y)$ reduces to a singleton. Therefore, by the corollary to [11, Theorem 2.2.4], we know that ϕ is continuously differentiable.

We now establish b). Let Y be an arbitrary stationary point of ψ , i.e., $\nabla\phi(Y) = 0$. This leads to

$$(\Gamma(Y) - E) \circ \Phi(Y) + (\mathcal{J}_Y F)^*((\Omega(Y) - E) \circ \Phi(Y)) = 0. \quad (38)$$

We only need to show $\Phi(Y) = 0$. By contradiction, we suppose that there exists an index (i, j) such that $\Phi_{ij}(Y) \neq 0$, which implies that one of the following results holds:

- 1). $Y_{ij} < 0$ and $F_{ij}(Y) = 0$.
- 2). $Y_{ij} = 0$ and $F_{ij}(Y) < 0$.
- 3). $Y_{ij} \neq 0$ and $F_{ij}(Y) \neq 0$.

For every case, we have

$$\left(\frac{Y_{ij}}{\|(Y_{ij}, F_{ij}(Y))\|} - 1 \right) \left(\frac{F_{ij}(Y)}{\|(Y_{ij}, F_{ij}(Y))\|} - 1 \right) > 0. \quad (39)$$

Notice that F is a linear operator. Then, for any $H \in \mathbb{R}^{n \times n}$, we have

$$F(Y + H) - F(Y) = \mathcal{J}_Y F(H).$$

By Proposition 3.16, F is monotone. Thus,

$$\begin{aligned} & \langle (\Omega(Y) - E) \circ \Phi(Y), (\mathcal{J}_Y F)^*((\Omega(Y) - E) \circ \Phi(Y)) \rangle \\ &= \langle (\Omega(Y) - E) \circ \Phi(Y), \mathcal{J}_Y F((\Omega(Y) - E) \circ \Phi(Y)) \rangle \geq 0. \end{aligned}$$

Hence, by using (38), we have

$$\langle (\Omega(Y) - E) \circ \Phi(Y), (\Gamma(Y) - E) \circ \Phi(Y) \rangle \leq 0.$$

By Proposition 3.17 and (39),

$$0 < \left(\frac{Y_{ij}}{\|(Y_{ij}, F_{ij}(Y))\|} - 1 \right) \left(\frac{F_{ij}(Y)}{\|(Y_{ij}, F_{ij}(Y))\|} - 1 \right) \Phi_{ij}^2(Y) \leq 0.$$

This is a contradiction. Therefore, $\Phi(Y) = 0$. □

We now establish the following useful lemma.

Lemma 3.20 *Let $Y \in \mathbb{R}^{n \times n}$ and $\mathcal{H}(Y) := \{D_1 \circ + D_2 \circ \mathcal{J}_Y F \mid -D_1, -D_2 \in \mathbb{R}_{++}^{n \times n}\}$. Then, every element in $\mathcal{H}(Y)$ is nonsingular.*

Proof: Let V be arbitrary element in $\mathcal{H}(Y)$. Then, there exist two matrices D_1, D_2 such that $-D_1, -D_2 \in \mathbb{R}_{++}^{n \times n}$ and

$$V = D_1 \circ + D_2 \circ \mathcal{J}_Y F. \quad (40)$$

For the sake of contradiction, we assume that V is singular. Then, there exists an $0 \neq H \in \mathbb{R}^{n \times n}$ such that $V(H) = 0$. By (40), we obtain $D_1 \circ H = -D_2 \circ \mathcal{J}_Y F(H)$ or

$$H_{ij} = -\frac{(D_2)_{ij}}{(D_1)_{ij}} (\mathcal{J}_Y F(H))_{ij} \quad \text{for } i, j = 1, \dots, n.$$

Since F is a linear operator, we obtain

$$F(Y + H) - F(Y) = \mathcal{J}_Y F(H). \quad (41)$$

By Proposition 3.16, F is monotone. Hence, F is also a P_0 function. The latter statement, together with (41) and $-D_1, -D_2 \in \mathbb{R}_{++}^{n \times n}$, implies that there exists an index (i_0, j_0) such that $H_{i_0 j_0} \neq 0$ and

$$H_{i_0 j_0} (\mathcal{J}_Y F(H))_{i_0 j_0} = (Y_{i_0 j_0} + H_{i_0 j_0} - Y_{i_0 j_0}) (F_{i_0 j_0}(Y + H) - F_{i_0 j_0}(Y)) \geq 0.$$

Hence,

$$0 < H_{i_0 j_0}^2 = -H_{i_0 j_0} \frac{(D_2)_{i_0 j_0}}{(D_1)_{i_0 j_0}} (\mathcal{J}_Y F(H))_{i_0 j_0} \leq 0,$$

i.e., $H_{i_0 j_0} = 0$. This is a contradiction because $H_{i_0 j_0} \neq 0$. Therefore, V is nonsingular. \square

We now construct a subset of $\partial_B \Phi(Y)$, which is easy to evaluate. Define

$$\mathcal{L}(Y) := \{L = S \circ + T \circ \mathcal{J}_Y F \mid (S, T) \in \mathcal{G}(Y)\}.$$

Here

$$\mathcal{G}(Y) := \{(S, T) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \mid (S, T) = (P(Y, Z), Q(Y, Z)), Z \in \mathcal{Z}(Y)\},$$

where $\mathcal{Z}(Y) := \{Z \in \mathbb{R}^{n \times n} \mid Z_{ij} \neq 0 \text{ if } (Y_{ij}, F_{ij}(Y)) \neq (0, 0)\}$ and $P(Y, Z)$ and $Q(Y, Z)$ are given by

$$\begin{aligned} & (P_{ij}(Y, Z), Q_{ij}(Y, Z)) \\ := & \begin{cases} \left(\frac{Z_{ij}}{\sqrt{Z_{ij}^2 + (\mathcal{J}_Y F(Z))_{ij}^2}} - 1, \frac{(\mathcal{J}_Y F(Z))_{ij}}{\sqrt{Z_{ij}^2 + (\mathcal{J}_Y F(Z))_{ij}^2}} - 1 \right), & \text{if } Y_{ij} = 0 = F_{ij}(Y), \\ \left(\frac{Y_{ij}}{\sqrt{Y_{ij}^2 + F_{ij}^2(Y)}} - 1, \frac{F_{ij}(Y)}{\sqrt{Y_{ij}^2 + F_{ij}^2(Y)}} - 1 \right), & \text{otherwise.} \end{cases} \end{aligned} \quad (42)$$

We state and prove the following result on $\mathcal{L}(Y)$.

Proposition 3.21 *For any $H \in \mathbb{R}^{n \times n}$,*

$$\mathcal{L}(Y)H \subseteq \partial_B \Phi(Y)H.$$

Proof: Let $L(H) = P(Y, Z) \circ H + Q(Y, Z) \circ \mathcal{J}_Y F(H) \in \mathcal{L}(Y)H$, where $Z \in \mathcal{Z}(Y)$ and $P(Y, Z)$ and $Q(Y, Z)$ are defined in (42). Let $Y^k := Y + \epsilon^k Z$, where $0 < \epsilon^k \rightarrow 0$. Notice that for $Y_{ij} = 0 = F_{ij}(Y)$, $Z_{ij} \neq 0$. By continuity, we can assume that ϵ^k is sufficiently small so that $Y_{ij} \neq 0$ or $F_{ij}(Y) \neq 0$ for all i, j and Φ is Fréchet differentiable at Y^k .

In the following, we show that $\lim_{k \rightarrow \infty} \mathcal{J}_{Y^k} \Phi(H) = L(H)$. If $(Y_{ij}, F_{ij}(Y)) \neq (0, 0)$, then by continuity it is easy to see that $\lim_{k \rightarrow \infty} (\mathcal{J}_{Y^k} \Phi(H))_{ij} = (L(H))_{ij}$. If $Y_{ij} = 0 = F_{ij}(Y)$, then Proposition 3.17 implies

$$(\mathcal{J}_{Y^k} \Phi(H))_{ij} = \left(\frac{\epsilon^k Z_{ij}}{\sqrt{(\epsilon^k)^2 Z_{ij}^2 + F_{ij}^2(Y^k)}} - 1 \right) + \left(\frac{F_{ij}(Y^k)}{\sqrt{(\epsilon^k)^2 Z_{ij}^2 + F_{ij}^2(Y^k)}} - 1 \right) (\mathcal{J}_{Y^k} F(H))_{ij}. \quad (43)$$

Since F is a linear operator, we have, for $Y_{ij} = 0 = F_{ij}(Y)$,

$$F_{ij}(Y^k) = F_{ij}(Y) + \epsilon^k (\mathcal{J}_{Y^k} F(Z))_{ij} = \epsilon^k (\mathcal{J}_{Y^k} F(Z))_{ij}. \quad (44)$$

By (43), (44), and the continuity of $\mathcal{J}_{Y^k} F$, we obtain $\lim_{k \rightarrow \infty} (\mathcal{J}_{Y^k} \Phi(H))_{ij} = (L(H))_{ij}$ if $Y_{ij} = 0 = F_{ij}(Y)$. \square

Now, we propose a globalized Newton-type method for solving (36). From Proposition 3.21, it seems natural to solve the linear system

$$L(\Delta Y^k) = -\Phi(Y^k), \quad L \in \mathcal{L}(Y^k).$$

However, we note that the nonsingularity of an element $L \in \mathcal{L}(Y^k)$ is not guaranteed. We may modify L by using the similar technique in [48]. Define

$$\begin{aligned} \tilde{\mathcal{L}}(Y) &:= \{ \tilde{L} = (S + \Delta S) \circ +(T + \Delta T) \circ \mathcal{J}_Y F \mid (S, T) \in \mathcal{G}(Y), \\ &\quad ((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathcal{Q}(Y, S_{ij}, T_{ij}) \text{ for } i, j = 1, \dots, n \}, \end{aligned} \quad (45)$$

where

$$\mathcal{Q}(Y, S_{ij}, T_{ij}) := \begin{cases} \left\{ ((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathbb{R}^2 \mid (\Delta S)_{ij} = \frac{\theta(\phi(Y))}{T_{ij}}, (\Delta T)_{ij} = 0 \right\}, & \text{if } -\delta < S_{ij} \text{ and } T_{ij} \leq -\delta, \\ \left\{ ((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathbb{R}^2 \mid (\Delta S)_{ij} = \mu \frac{\theta(\phi(Y))}{T_{ij}}, (\Delta T)_{ij} = (1 - \mu) \frac{\theta(\phi(Y))}{S_{ij}}, \mu \in [0, 1] \right\}, \\ & \text{if } S_{ij} \leq -\delta \text{ and } T_{ij} \leq -\delta, \\ \left\{ ((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathbb{R}^2 \mid (\Delta S)_{ij} = 0, (\Delta T)_{ij} = \frac{\theta(\phi(Y))}{S_{ij}} \right\}, & \text{if } S_{ij} \leq -\delta \text{ and } -\delta < T_{ij}. \end{cases} \quad (46)$$

Here, $\delta \in (0, 1 - 1/\sqrt{2})$ and $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a nondecreasing continuous function such that $\theta(0) = 0$ and $\theta(t) > 0$ for all $t > 0$. If $\phi(Y) > 0$, then it is easy to obtain that $-(S + \Delta S), -(T + \Delta T) \in \mathbb{R}_{++}^{n \times n}$. In this case, by Lemma 3.20, we know that any element $\tilde{L} \in \mathcal{L}(Y)$ is nonsingular.

Proposition 3.22 *Let $L = S \circ +T \circ \mathcal{J}_Y F \in \mathcal{L}(Y)$ and $\tilde{L} = L + \Delta S \circ +\Delta T \circ \mathcal{J}_Y F \in \tilde{\mathcal{L}}(Y)$, where $(S, T) \in \mathcal{G}(Y)$ and $((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathcal{Q}(Y, S_{ij}, T_{ij})$ for all i, j . Then $(\tilde{L} - L)^* L$ is positive definite.*

Proof: Notice that $\tilde{L} - L = \Delta S \circ + \Delta T \circ \mathcal{J}_Y F$. Then

$$\begin{aligned} (\tilde{L} - L)^* L &= (\Delta S \circ + \Delta T \circ \mathcal{J}_Y F)^*(S \circ + T \circ \mathcal{J}_Y F \in \mathcal{L}(Y)) \\ &= \Delta S \circ S \circ + \Delta S \circ T \circ \mathcal{J}_Y F + (\mathcal{J}_Y F)^* \Delta T \circ S \circ + (\mathcal{J}_Y F)^* \Delta T \circ T \circ \mathcal{J}_Y F. \end{aligned}$$

Since $S_{ij}, T_{ij}, (\Delta S)_{ij}, (\Delta T)_{ij} \leq 0$ for all i, j , both $\Delta S \circ S \circ$ and $\mathcal{J}_Y F^* \Delta T \circ T \circ \mathcal{J}_Y F$ are positive semidefinite.

Next, we establish the positive semidefiniteness of $(\Delta S \circ T \circ \mathcal{J}_Y F + (\mathcal{J}_Y F)^* \Delta T \circ S \circ)$. For any $H \in \mathbb{R}^{n \times n}$,

$$\begin{aligned} &\langle (\Delta S \circ T \circ \mathcal{J}_Y F + (\mathcal{J}_Y F)^* \Delta T \circ S \circ) H, H \rangle \\ &= \langle \Delta S \circ T \circ \mathcal{J}_Y F(H), H \rangle + \langle (\mathcal{J}_Y F)^* \Delta T \circ S \circ H, H \rangle \\ &= \langle \Delta S \circ T \circ \mathcal{J}_Y F(H), H \rangle + \langle H, \Delta T \circ S \circ \mathcal{J}_Y F(H) \rangle \\ &= \langle H, (\Delta S \circ T \circ + \Delta T \circ S \circ)(\mathcal{J}_Y F(H)) \rangle. \end{aligned}$$

Thus, we only need to show the positive semidefiniteness of $(\Delta S \circ T \circ + \Delta T \circ S \circ) \mathcal{J}_Y F$. Since $\mathcal{J}_Y F$ is positive semidefinite, it is enough to show that $(\Delta S \circ T \circ + \Delta T \circ S \circ)$ is positive semidefinite. Notice that $((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathcal{Q}(Y, S_{ij}, T_{ij})$ for all i, j . This implies that for all i, j ,

$$((\Delta S)_{ij} T_{ij}, (\Delta T)_{ij} S_{ij}) = \begin{cases} (\theta(\phi(Y)), 0), & \text{if } -\delta < S_{ij} \text{ and } T_{ij} \leq -\delta, \\ (\mu\theta(\phi(Y)), (1 - \mu)\theta(\phi(Y))), & \text{if } S_{ij} \leq -\delta \text{ and } T_{ij} \leq -\delta, \\ (0, \theta(\phi(Y))), & \text{if } S_{ij} \leq -\delta \text{ and } -\delta < T_{ij}. \end{cases}$$

Thus, $(\Delta S)_{ij} T_{ij} + (\Delta T)_{ij} S_{ij} = \theta(\phi(Y))$ for all i, j , where $\theta(\psi(Y)) \geq 0$. Therefore, $\Delta S \circ T \circ + \Delta T \circ S \circ = \theta(\phi(Y)) E \circ$ is positive semidefinite. This completes the proof. \square

We now state the Newton-type algorithm for solving Problem (30) as follows.

Algorithm 3.23 (A Newton-Type Method for the NIEP)

Step 0. Give $Y^0 \in \mathbb{R}^{n \times n}$, $\eta \in (0, 1)$, and $\rho, \sigma \in (0, 1/2)$. $k := 0$.

Step 1. If $\phi(Y^k) = 0$, then stop. Otherwise, go to Step 2.

Step 2. Select an element $\tilde{L}_k \in \tilde{\mathcal{L}}(Y^k)$. Apply an iterative method (e.g., the transpose-free quasi-minimal residual (TFQMR) method [20]) to solving

$$\tilde{L}_k(\Delta Y^k) + \Phi(Y^k) = 0 \quad (47)$$

for $\Delta Y^k \in \mathbb{R}^{n \times n}$ such that

$$\|\tilde{L}_k(\Delta Y^k) + \Phi(Y^k)\| \leq \eta_k \|\Phi(Y^k)\|, \quad (48)$$

and

$$\langle \nabla \phi(Y^k), \Delta Y^k \rangle \leq -\eta_k \langle \Delta Y^k, \Delta Y^k \rangle, \quad (49)$$

where $\eta_k := \min\{\eta, \|\Phi(Y^k)\|\}$. If (48) and (49) are not attainable, then let

$$\Delta Y^k := -\nabla \phi(Y^k).$$

Step 3. Let l_k be the smallest nonnegative integer l such that

$$\phi(Y^k + \rho^l \Delta Y^k) - \phi(Y^k) \leq \sigma \rho^l \langle \nabla \phi(Y^k), \Delta Y^k \rangle.$$

Set

$$Y^{k+1} := Y^k + \rho^{l_k} \Delta Y^k.$$

Step 4. Replace k by $k + 1$ and go to Step 1.

We note that, in Steps 2 and 3 of Algorithm 3.23, we need to compute $\nabla \phi(Y^k)$ at the k th iteration. By Propositions 3.19 and 3.21, it is easy to see that $\nabla \phi(Y^k) = L_k^* \Phi(Y^k)$, where $L_k \in \mathcal{L}(Y^k)$. Since the problem size is n^2 , the direct solution of (47) needs $O(n^4)$ operations, which is very costly if the problem size is large. Therefore, in Algorithm 3.23, we propose to solve (47) inexactly by iterative methods. Moreover, we will see in Subsection 3.3 that requirement (49) is reasonable (see Proposition 3.24 below).

3.3 Convergence Analysis

In this subsection, we shall establish the global and quadratic convergence of Algorithm 3.23. First, we state the following result on the descent property of the solution to $\tilde{L}(\Delta Y) + \Phi(Y) = 0$, where $\tilde{L} \in \tilde{\mathcal{L}}(Y)$.

Proposition 3.24 *If $Y \in \mathbb{R}^{n \times n}$ is not a solution of Problem (30), then any solution $\Delta Y \in \mathbb{R}^{n \times n}$ to*

$$\tilde{L}(\Delta Y) + \Phi(Y) = 0, \quad \tilde{L} \in \tilde{\mathcal{L}}(Y) \tag{50}$$

is a descent direction of ϕ , i.e., $\langle \nabla \phi(Y), \Delta Y \rangle < 0$.

Proof: We first show $\langle \nabla \phi(Y), \Delta Y \rangle \leq 0$. Let $\tilde{L} = L + \Delta S \circ + \Delta T \circ \mathcal{J}_Y F \in \tilde{\mathcal{L}}(Y)$, where $L \in \mathcal{L}(Y)$ and $((\Delta S)_{ij}, (\Delta T)_{ij}) \in \mathcal{Q}(Y, S_{ij}, T_{ij})$ for all i, j . By Proposition 3.19 a) and Proposition 3.21, we have

$$\begin{aligned} \langle \nabla \phi(Y), \Delta Y \rangle &= \langle L^* \Phi(Y), \Delta Y \rangle = -\langle L^* \tilde{L}(\Delta Y), \Delta Y \rangle = -\langle \Delta Y, \tilde{L}^* L \Delta Y \rangle \\ &= -\langle \Delta Y, L^* L \Delta Y \rangle - \langle \Delta Y, (\tilde{L} - L)^* L \Delta Y \rangle. \end{aligned} \tag{51}$$

By Proposition 3.22, we deduce that $\langle \nabla \phi(Y), \Delta Y \rangle \leq 0$.

Next, we prove that if $\langle \nabla \phi(Y), \Delta Y \rangle = 0$ holds for a solution ΔY to (50), then Y is a solution of Problem (30). By contradiction, we assume that $\langle \nabla \phi(Y), \Delta Y \rangle = 0$ for a solution ΔY to (50), but Y is not a solution of Problem (30). Since $\Phi(Y) \neq 0$, we have $\Delta Y \neq 0$. Let $\tilde{L} = (S + \Delta S) \circ + (T + \Delta T) \circ \mathcal{J}_Y F \in \tilde{\mathcal{L}}(Y)$ be such that $\tilde{L}(\Delta Y) + \Phi(Y) = 0$. Define $\hat{L} := \hat{S} \circ + \hat{T} \circ \mathcal{J}_Y F$, where for all i, j ,

$$(\hat{S}_{ij}, \hat{T}_{ij}) = \begin{cases} (S_{ij} + (\Delta S)_{ij}, T_{ij} + (\Delta T)_{ij}), & \text{if } \omega(Y_{ij}, F_{ij}(Y)) = 0, \\ (S_{ij}, T_{ij}), & \text{if } \omega(Y_{ij}, F_{ij}(Y)) \neq 0. \end{cases}$$

By Lemma 3.20, \hat{L} is nonsingular. Thus, for all i, j ,

$$(\hat{L} \Delta Y)_{ij} = \begin{cases} (\tilde{L} \Delta Y)_{ij} = -(\Phi(Y))_{ij} = 0, & \text{if } \omega(Y_{ij}, F_{ij}(Y)) = 0, \\ (L \Delta Y)_{ij}, & \text{if } \omega(Y_{ij}, F_{ij}(Y)) \neq 0. \end{cases}$$

By (51) and the positive semidefiniteness of $(\tilde{L} - L)^*L$, it follows from $\langle \nabla\phi(Y), \Delta Y \rangle = 0$ that $L\Delta Y = 0$. Thus, if $\phi(Y_{ij}, F_{ij}(Y)) \neq 0$, $(\widehat{L}\Delta Y)_{ij} = (L\Delta Y)_{ij} = 0$. Hence, $\widehat{L}\Delta Y = 0$ and, taking into account the nonsingularity of \widehat{L} , we deduce $\Delta Y = 0$. This is a contradiction since $\Delta Y \neq 0$ and the proof is complete. \square

By using the similar proof of Theorem 11 (a) in [13] or Theorem 3.1 in [48], we can derive the following theorem on the global convergence of Algorithm 3.23. We omit it here.

Theorem 3.25 *Any accumulation point of the sequence $\{Y^k\}$ generated by Algorithm 3.23 is a solution to Problem (30).*

The remaining part of this subsection is concerned with the quadratic convergence of Algorithm 3.23. Here, we assume that every element in $\partial_B\Phi(\cdot)$ is nonsingular (the nonsingularity of $\partial_B\Phi(\cdot)$ is studied in Subsection 3.4).

Assumption 3.26 *All elements in $\partial_B\Phi(\bar{Y})$ are nonsingular, where \bar{Y} is an accumulation point of the sequence $\{Y^k\}$ generated by Algorithm 3.23.*

Under Assumption 3.26, we obtain the following lemma. The proof is similar to that of Theorem 11 (b),(c) in [13] and we omit it.

Lemma 3.27 *Let \bar{Y} be an accumulation point of the sequence $\{Y^k\}$ generated by Algorithm 3.23. Suppose that Assumption 3.26 is satisfied. Then the sequence $\{Y^k\}$ converges to \bar{Y} .*

We now establish the quadratic convergence of Algorithm 3.23.

Theorem 3.28 *Let \bar{Y} be an accumulation point of the sequence $\{Y^k\}$ generated by Algorithm 3.23. Suppose that Assumption 3.26 is satisfied and the function θ is such that $\theta(t) = O(\sqrt{t})$. Then the sequence $\{Y^k\}$ converges to \bar{Y} quadratically.*

Proof: By Proposition 3.18, Φ is strongly semismooth. Thus, by Proposition 3.21, we get

$$\|\Phi(Y^k) - \Phi(\bar{Y}) - L_k(Y^k - \bar{Y})\| = O(\|Y^k - \bar{Y}\|^2). \quad (52)$$

Moreover, by the assumption that $\theta(t) = O(\sqrt{t})$, we have, for all k sufficiently large,

$$\theta(\phi(Y^k)) = O\left(\sqrt{\phi(Y^k)}\right) = O(\|\Phi(Y^k)\|) = O(\|\Phi(Y^k) - \Phi(\bar{Y})\|) = O(\|Y^k - \bar{Y}\|). \quad (53)$$

Since Assumption 3.26 is satisfied, by Lemma 3.27, $\lim_{k \rightarrow \infty} Y^k = \bar{Y}$. Hence, by Proposition 3.21, for all k sufficiently large, L_k is nonsingular and $\|L_k^{-1}\|$ is uniformly bounded. Since $\theta(\phi(Y^k)) \rightarrow 0$, it is easy to know from (45) and (46) that for all k sufficiently large, $\|\tilde{L}_k - L_k\| = O(\theta(\phi(Y^k)))$ are small enough and then $\|\tilde{L}_k^{-1}\|$ is uniformly bounded. Thus, for all k sufficiently large, an iterative method can find ΔY^k such that both (48) and (49) are satisfied. Therefore, by (48),

(52), and (53), for all k sufficiently large,

$$\begin{aligned}
& \|Y^k + \Delta Y^k - \bar{Y}\| \\
&= \|\tilde{L}_k^{-1}(\Phi(\bar{Y}) - \Phi(Y^k)) + L_k(Y^k - \bar{Y}) + (\tilde{L}_k - L_k)(Y^k - \bar{Y}) + \tilde{L}_k \Delta Y^k + \Phi(Y^k)\| \\
&\leq \|\tilde{L}_k^{-1}\|\|\Phi(\bar{Y}) - \Phi(Y^k) + L_k(Y^k - \bar{Y})\| + \|\tilde{L}_k^{-1}\|\|\tilde{L}_k - L_k\|\|Y^k - \bar{Y}\| \\
&\quad + \|\tilde{L}_k^{-1}\|\|\tilde{L}_k \Delta Y^k + \Phi(Y^k)\| \\
&= O(\|Y^k - \bar{Y}\|^2) + O(\|Y^k - \bar{Y}\|^2) + \eta_k \|\tilde{L}_k^{-1}\|\|\Phi(Y^k)\| \\
&= O(\|Y^k - \bar{Y}\|^2) + O(\|\Phi(Y^k)\|^2) \\
&= O(\|Y^k - \bar{Y}\|^2).
\end{aligned} \tag{54}$$

On the other hand, for any $\tau > 0$ and for all k sufficiently large,

$$\begin{aligned}
& -\langle \nabla \phi(Y^k), \Delta Y^k \rangle \\
&= -\langle L_k^* \Phi(Y^k), \Delta Y^k \rangle \\
&= \langle (\tilde{L}_k - L_k - \tilde{L}_k)^* \Phi(Y^k), \tilde{L}_k^{-1}(\tilde{L}_k \Delta Y^k + \Phi(Y^k) - \Phi(Y^k)) \rangle \\
&= \langle \Phi(Y^k), (\tilde{L}_k - L_k) \tilde{L}_k^{-1}(\tilde{L}_k \Delta Y^k + \Phi(Y^k)) \rangle - \langle \Phi(Y^k), (\tilde{L}_k - L_k) \tilde{L}_k^{-1} \Phi(Y^k) \rangle \\
&\quad - \langle \Phi(Y^k), \tilde{L}_k \Delta Y^k + \Phi(Y^k) \rangle + \langle \Phi(Y^k), \Phi(Y^k) \rangle \\
&\leq \|\tilde{L}_k - L_k\|\|\tilde{L}_k^{-1}\|\|\Phi(Y^k)\|\|\tilde{L}_k \Delta Y^k + \Phi(Y^k)\| + \|\tilde{L}_k - L_k\|\|\tilde{L}_k^{-1}\|\|\Phi(Y^k)\|^2 \\
&\quad + \|\Phi(Y^k)\|\|\tilde{L}_k \Delta Y^k + \Phi(Y^k)\| + \|\Phi(Y^k)\|^2 \\
&\leq 2(1 + \eta_k + (\eta_k + 1)\|\tilde{L}_k - L_k\|\|\tilde{L}_k^{-1}\|)\phi(Y^k) \\
&\leq 2(1 + \tau)\phi(Y^k),
\end{aligned} \tag{55}$$

where the last step uses the facts that $\|\tilde{L}_k^{-1}\|$ is uniformly bounded and η_k and $\|\tilde{L}_k - L_k\|$ are small enough for all k sufficiently large. Moreover, there exists a constant $\xi > 0$ such that

$$\begin{aligned}
\|Y^k - \bar{Y}\| &\leq \|Y^k + \Delta Y^k - \bar{Y}\| + \|\Delta Y^k\| \\
&\leq \|Y^k + \Delta Y^k - \bar{Y}\| + \|\tilde{L}_k^{-1}\|\|\tilde{L}_k \Delta Y^k + \Phi(Y^k)\| + \|\tilde{L}_k^{-1}\|\|\Phi(Y^k)\| \\
&\leq \|Y^k + \Delta Y^k - \bar{Y}\| + (1 + \eta_k)\|\tilde{L}_k^{-1}\|\|\Phi(Y^k)\| \\
&\leq \|Y^k + \Delta Y^k - \bar{Y}\| + \xi\|\Phi(Y^k)\|
\end{aligned} \tag{56}$$

for all k sufficiently large. By (54), for all k large enough, there exists a sufficiently small $\varsigma > 0$ such that $\|Y^k + \Delta Y^k - \bar{Y}\| \leq \varsigma\|Y^k - \bar{Y}\|$. This, together with (56), leads to

$$\|Y^k + \Delta Y^k - \bar{Y}\| \leq \frac{\varsigma \xi}{1 - \varsigma} \|\Phi(Y^k)\|. \tag{57}$$

Since Φ is Lipschitz continuous near \bar{Y} with a Lipschitz constant $\varrho > 0$, we have by (57), for all k sufficiently large,

$$\|\Phi(Y^k + \Delta Y^k)\| = \|\Phi(Y^k + \Delta Y^k) - \Phi(\bar{Y})\| \leq \varrho \|Y^k + \Delta Y^k - \bar{Y}\| \leq \varrho \frac{\varsigma \xi}{1 - \varsigma} \|\Phi(Y^k)\|$$

or

$$\phi(Y^k + \Delta Y^k) \leq \left(\frac{\varsigma \xi}{1 - \varsigma} \right)^2 \phi(Y^k), \quad \zeta := \varrho \xi. \quad (58)$$

Since $\sigma \in (0, 1/2)$, by (55), we can choose a sufficiently small $\tau > 0$ for which $2(1 + \tau)\sigma < 1$. Also, by (58), $\varsigma > 0$ can be taken small enough such that

$$1 - \left(\frac{\varsigma \xi}{1 - \varsigma} \right)^2 > 2(1 + \tau)\sigma.$$

From (55) and (58), for all k sufficiently large, we obtain

$$\begin{aligned} \phi(Y^k + \Delta Y^k) - \phi(Y^k) &\leq - \left(1 - \left(\frac{\varsigma \xi}{1 - \varsigma} \right)^2 \right) \phi(Y^k) \\ &\leq -2(1 + \tau)\sigma \phi(Y^k) \leq \sigma \langle \nabla \phi(Y^k), \Delta Y^k \rangle, \end{aligned}$$

which implies that for all k sufficiently large,

$$Y^{k+1} = Y^k + \Delta Y^k.$$

Thus, by exploiting (54), the proof is complete. \square

3.4 Nonsingularity Conditions of $\partial\Phi(\cdot)$

Finally, we discuss the nonsingularity of $\partial\Phi(\cdot)$ at a solution \bar{Y} of Problem (30). Notice that the nonsmooth equation (36) is equivalent to the nonsmooth equation (32). As in Subsection 3.1, both $\partial H(\cdot, \cdot)$ and $\partial\Phi(\cdot)$ should share the similar nonsingularity conditions. Let (\bar{Y}, \bar{Z}) be the KKT point of Problem (30). The critical cone $\mathcal{C}(\bar{Y})$ of Problem (30) at \bar{Y} is defined by

$$\mathcal{C}(\bar{Y}) := \{d \mid d \in \mathcal{T}_{\mathbb{R}_+^{n \times n}}(\bar{Y}), \mathcal{J}_Y f(\bar{Y})d \leq 0\}.$$

Since \bar{Y} is a stationary point of Problem (30), we have

$$\mathcal{C}(\bar{Y}) = \{d \mid d \in \mathcal{T}_{\mathbb{R}_+^{n \times n}}(\bar{Y}), \mathcal{J}_Y f(\bar{Y})d = 0\}.$$

Since (\bar{Y}, \bar{Z}) is a KKT point of Problem (30), we have by [14],

$$-\bar{Z} \in \mathcal{N}_{\mathbb{R}_+^{n \times n}}(\bar{Y}) \iff \mathbb{R}_+^{n \times n} \ni \bar{Z} \perp \bar{Y} \in \mathbb{R}_+^{n \times n}.$$

Then $C := \bar{Y} - \bar{Z}$ can be rewritten in analogy to (4), i.e.,

$$\bar{Y} = P(C_{\alpha \cup \beta}) \quad \text{and} \quad \bar{Z} = -P(C_{\beta \cup \gamma}).$$

By (10) and (11), we get

$$\begin{aligned}\mathcal{T}_{\mathbb{R}_+^{n \times n}}(\bar{Y}) &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\}, \\ \mathcal{T}_{\mathbb{R}_+^{n \times n}}(\bar{Y}) \cap (-\bar{Z})^\perp &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta \geq 0, H_\gamma = 0\}, \\ \text{lin}(\mathcal{T}_{\mathbb{R}_+^{n \times n}}(\bar{Y})) &= \{H \in \mathbb{R}^{n \times n} \mid H_\beta = 0, H_\gamma = 0\}.\end{aligned}$$

Since $\mathcal{M}(\bar{Y})$ is nonempty, by (13), we infer that

$$\mathcal{C}(\bar{Y}) = \{d \mid d \in \mathcal{C}(C; \mathbb{R}_+^{n \times n})\} = \mathcal{C}(C; \mathbb{R}_+^{n \times n}).$$

Thus

$$\text{aff}(\mathcal{C}(\bar{Y})) = \{d \in \mathbb{R}^{n \times n} \mid d_\gamma = 0\}.$$

Therefore, based on Theorem 3.14, we can prove the following result on the nonsingularity of $\partial\Phi(\cdot)$.

Theorem 3.29 *Let \bar{Y} be an accumulation point of the sequence $\{Y^k\}$ generated by Algorithm 3.23. Let $C := \bar{Y} - \bar{Z}$, where $\bar{Z} := F(\bar{Y})$. Suppose that the strong second order sufficient condition holds at \bar{Y} i.e.,*

$$\langle d, \mathcal{J}_{\bar{Y}}^2 f(\bar{Y})d \rangle > 0 \quad \forall d \in \text{aff}(\mathcal{C}(\bar{Y})) \setminus \{0\}. \quad (59)$$

Then any element in $\partial\Phi(\bar{Y})$ is nonsingular.

Proof: By Theorem 3.25, we know \bar{Y} is solution to Problem (30). Let V be an arbitrary element in $\partial\Phi(\bar{Y})$. Let $\Delta Y \in \mathbb{R}^{n \times n}$ be such that

$$V(\Delta Y) = 0.$$

By Proposition 3.17, we have

$$V(\Delta Y) = (\Gamma(\bar{Y}) - E) \circ \Delta Y + (\Omega(\bar{Y}) - E) \circ (\mathcal{J}_{\bar{Y}}^2 f(\bar{Y})(\Delta Y)) = 0. \quad (60)$$

By the assumption that $C := \bar{Y} - \bar{Z}$ and $\bar{Z} := F(\bar{Y})$, we obtain

$$\begin{cases} \alpha & := \{(i, j) \mid C_{ij} > 0\} = \{(i, j) \mid \bar{Y}_{ij} > 0, F_{ij}(\bar{Y}) = 0\}, \\ \beta & := \{(i, j) \mid C_{ij} = 0\} = \{(i, j) \mid \bar{Y}_{ij} = 0, F_{ij}(\bar{Y}) = 0\}, \\ \gamma & := \{(i, j) \mid C_{ij} < 0\} = \{(i, j) \mid \bar{Y}_{ij} = 0, F_{ij}(\bar{Y}) > 0\}, \end{cases}$$

which implies that

$$\Gamma_{ij}(\bar{Y}) = \begin{cases} 1, & \text{if } (i, j) \in \alpha, \\ \Gamma_{ij}, & \text{if } (i, j) \in \beta, \\ 0, & \text{if } (i, j) \in \gamma, \end{cases} \quad \text{and} \quad \Omega_{ij}(\bar{Y}) = \begin{cases} 0, & \text{if } (i, j) \in \alpha, \\ \Omega_{ij}, & \text{if } (i, j) \in \beta, \\ 1, & \text{if } (i, j) \in \gamma, \end{cases} \quad (61)$$

where $\|(\Gamma_{ij}, \Omega_{ij})\| \leq 1$. From (61), it follows that (60) is reduced to

$$\begin{cases} (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_\alpha = 0, \\ (\Gamma_\beta(\bar{Y}) - E_\beta) \circ \Delta Y_\beta + (\Omega_\beta(\bar{Y}) - E_\beta) \circ (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_\beta = 0, \\ \Delta Y_\gamma = 0. \end{cases} \quad (62)$$

Let

$$\begin{aligned} \beta_1 &:= \{(i, j) \in \beta \mid \Gamma_{ij} = 1, \Omega_{ij} = 0\}, \\ \beta_2 &:= \{(i, j) \in \beta \mid \Gamma_{ij} = 0, \Omega_{ij} = 1\}, \\ \beta_3 &:= \beta \setminus (\beta_1 \cup \beta_2). \end{aligned}$$

Then (62) takes the form of

$$\begin{cases} (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{\alpha \cup \beta_1} = 0, \\ (\Gamma_{\beta_3}(\bar{Y}) - E_{\beta_3}) \circ \Delta Y_{\beta_3} + (\Omega_{\beta_3}(\bar{Y}) - E_{\beta_3}) \circ (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{\beta_3} = 0, \\ \Delta Y_{\gamma \cup \beta_2} = 0, \end{cases} \quad (63)$$

where $-(\Gamma_{\beta_3}(\bar{Y}) - E_{\beta_3}), -(\Omega_{\beta_3}(\bar{Y}) - E_{\beta_3}) \in \mathbb{R}_{++}^{|\beta_3| \times |\beta_3|}$. By the assumption that the strong second order sufficient condition (59) holds at \bar{Y} , we have

$$\begin{aligned} \langle \Delta Y_{\alpha \cup \beta_1}, (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{\alpha \cup \beta_1} \rangle &> 0 \quad \forall \Delta Y_{\alpha \cup \beta_1} \neq 0, \\ \langle \Delta Y_{\beta_3}, (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{\beta_3} \rangle &> 0 \quad \forall \Delta Y_{\beta_3} \neq 0. \end{aligned} \quad (64)$$

From the first equality of (63) and the first inequality of (64), we get

$$\Delta Y_{\alpha \cup \beta_1} = 0. \quad (65)$$

From the second equality of (63), we deduce that

$$(\Delta Y_{\beta_3})_{ij} = -\frac{\Omega_{ij}(\bar{Y}) - E_{ij}}{\Gamma_{ij}(\bar{Y}) - E_{ij}} (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{ij}, \quad \forall (i, j) \in \beta_3.$$

This leads to

$$\langle \Delta Y_{\beta_3}, (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{\beta_3} \rangle = -\sum_{(i,j) \in \beta_3} \frac{\Omega_{ij}(\bar{Y}) - E_{ij}}{\Gamma_{ij}(\bar{Y}) - E_{ij}} (\mathcal{J}_{Y^2}^2 f(\bar{Y})(\Delta Y))_{ij}^2 \leq 0,$$

which, together with the second inequality of (64), implies that

$$\Delta Y_{\beta_3} = 0. \quad (66)$$

Therefore, by (65), (66), and the last equality of (63), we know that V is nonsingular. \square

Remark 3.30 We assume that X is full column rank since X is the real matrix of partial prescribed eigenvectors. Let the QR decomposition of X be given by

$$X = U \begin{bmatrix} R \\ 0 \end{bmatrix},$$

where $U = [U_1, U_2]$ is an $n \times n$ orthogonal matrix with $U_1 \in \mathbb{R}^{n \times p}$ and R is a $p \times p$ nonsingular upper triangular matrix. If

$$U_1^T d \neq 0 \quad \forall d \in \text{aff}(\mathcal{C}(\bar{Y})) \setminus \{0\},$$

then we have

$$\langle d, \mathcal{J}_{\bar{Y}}^2 f(\bar{Y})d \rangle = \langle d, XX^T d \rangle = \langle X^T d, X^T d \rangle = \langle R^T(U_1 d), R^T(U_1 d) \rangle > 0 \quad \forall d \in \text{aff}(\mathcal{C}(\bar{Y})) \setminus \{0\}.$$

In this case, the strong second order sufficient condition (59) holds at \bar{Y} .

4 Extensions

In this section, we extend the proposed Newton-type method to the symmetric nonnegative inverse problem and to the cases of lower bounds and of prescribed entries.

4.1 The Symmetric Nonnegative Inverse Eigenvalue Problem with Partial Eigendata

The symmetric nonnegative inverse problem with partial eigendata can be stated as follows.

SNIEP. Construct a nontrivial n -by- n symmetric nonnegative matrix A from a set of measured partial eigendata $\{(\lambda_k, \mathbf{x}_k)\}_{k=1}^p$ ($p \leq n$).

Let $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^{p \times p}$ and $X = [x_1, \dots, x_p] \in \mathbb{R}^{n \times p}$. Then the SNIEP amounts to determining a matrix $A \in \mathbb{SR}_+^{n \times n}$ such that

$$AX = X\Lambda. \tag{67}$$

Let $K = X^T$ and $B = (X\Lambda)^T$. We note that $\bar{A} \in \mathbb{SR}_+^{n \times n}$ is a solution to (67) if and only if $\bar{Y} = \bar{A} \in \mathbb{SR}^{n \times n}$ is a global solution of the following convex optimization problem

$$\begin{cases} \min & f(Y) := \frac{1}{2} \|KY - B\|^2 \\ \text{s.t.} & Y \in \mathbb{SR}_+^{n \times n} \end{cases} \tag{68}$$

with $f(\bar{Y}) = 0$. It is well-known that the KKT condition for Problem (68) is given by

$$\mathcal{J}_Y f(Y) - Z = 0, \quad Y \in \mathbb{SR}_+^{n \times n}, \quad Z \in \mathbb{SR}_+^{n \times n}, \quad \langle Y, Z \rangle = 0. \tag{69}$$

Let

$$F(Y) := \mathcal{J}_Y f(Y) = \frac{1}{2} (K^T(KY - B) + (KY - B)^T K).$$

Then, the KKT condition (69) becomes the following MCP

$$Y \in \mathbb{SR}_+^{n \times n}, \quad F(Y) \in \mathbb{SR}_+^{n \times n}, \quad \langle Y, F(Y) \rangle = 0. \tag{70}$$

By using the well-known Fischer function ω defined in (35), the MCP (70) aims to find a solution $Y \in \mathbb{SR}^{n \times n}$ to the following equation

$$\Phi(Y) = 0 \quad (71)$$

where $\Phi_{ij}(Y) := \omega(Y_{ij}, F_{ij}(Y))$ for $i, j = 1, \dots, n$. Also, define the merit function $\phi : \mathbb{SR}^{n \times n} \rightarrow \mathbb{R}$ by

$$\phi(Y) := \frac{1}{2} \|\Phi(Y)\|^2. \quad (72)$$

Thus, based on the functions Φ defined in (71) and ϕ in (72), one may solve Problem (68) by Algorithm 3.23, where \bar{Y} is a solution. Therefore, we get a solution \bar{A} to the SNIEP as follows.

$$\bar{A} = \bar{Y}.$$

The global and quadratic convergence of Algorithm 3.23 for solving Problem (68) can be established as in Section 3.

4.2 The Case of Lower Bounds

4.2.1 The Nonsymmetric Case

In many applications, each entry of A is required to be equal to or greater than a nonnegative number. The NIEP with lower bounds is to find a matrix $A \in \mathbb{R}^{n \times n}$ such that

$$AX = X\Lambda, \quad A \geq \underline{A}, \quad (73)$$

where $\underline{A} \in \mathbb{R}_+^{n \times n}$ is a prescribed matrix and $A \geq \underline{A}$ means that $A - \underline{A} \geq 0$. By renaming $A := A - \underline{A}$, Problem (73) becomes a new NIEP, which aims to find a matrix $A \in \mathbb{R}_+^{n \times n}$ such that

$$AX = X\Lambda - \underline{A}X. \quad (74)$$

Let $K = X^T$ and $B = (X\Lambda - \underline{A}X)^T$. Then, Problem (74) is equivalent to finding a global solution to the following optimization problem

$$\begin{cases} \min & f(Y) := \frac{1}{2} \|KY - B\|^2 \\ \text{s.t.} & Y \in \mathbb{R}_+^{n \times n} \end{cases} \quad (75)$$

in the sense that \bar{A} is a solution to Problem (74) if and only if $\bar{Y} = \bar{A}^T$ is a global solution to Problem (75), with $f(\bar{Y}) = 0$. The KKT condition for Problem (75) is given by

$$\mathcal{J}_Y f(Y) - Z = 0, \quad Y \in \mathbb{R}_+^{n \times n}, \quad Z \in \mathbb{R}_+^{n \times n}, \quad \langle Y, Z \rangle = 0. \quad (76)$$

Let

$$F(Y) := \mathcal{J}_Y f(Y) = K^T(KY - B).$$

Then, the KKT condition (76) can be rewritten as the following MCP

$$Y \in \mathbb{R}_+^{n \times n}, \quad F(Y) \in \mathbb{R}_+^{n \times n}, \quad \langle Y, F(Y) \rangle = 0. \quad (77)$$

By using the well-known Fischer function ω defined in (35), solving the MCP (77) is equivalent to the solution of the following nonsmooth equation

$$\Phi(Y) = 0, \quad (78)$$

where $\Phi_{ij}(Y) := \omega(Y_{ij}, F_{ij}(Y))$ for $i, j = 1, \dots, n$. Also, define the merit function $\phi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ by

$$\phi(Y) := \frac{1}{2} \|\Phi(Y)\|^2. \quad (79)$$

Hence, based on the functions Φ defined in (78) and ϕ in (79), one may apply Algorithm 3.23 to Problem (75), where \bar{Y} is a solution. Thus, we find a solution \bar{A} to the NIEP with lower bounds (73) as follows

$$\bar{A} = \bar{Y}^T + \underline{A}.$$

As in Section 3, we can show the global and quadratic convergence of Algorithm 3.23 for solving Problem (75).

4.2.2 The Symmetric Case

The SNIEP with lower bounds amounts to finding a matrix $A \in \mathbb{SR}^{n \times n}$ such that

$$AX = X\Lambda, \quad A \geq \underline{A}, \quad (80)$$

where $\underline{A} \in \mathbb{SR}_+^{n \times n}$ is a prescribed matrix. By renaming $A := A - \underline{A}$, Problem (80) becomes a new SNIEP, which is to determine a matrix $A \in \mathbb{SR}_+^{n \times n}$ such that

$$AX = X\Lambda - \underline{A}X. \quad (81)$$

Let $K = X^T$ and $B = (X\Lambda - \underline{A}X)^T$. Then, Problem (81) is equivalent to finding a global solution to the following optimization problem

$$\begin{cases} \min & f(Y) := \frac{1}{2} \|KY - B\|^2 \\ \text{s.t.} & Y \in \mathbb{SR}_+^{n \times n} \end{cases} \quad (82)$$

in the sense that \bar{A} is a solution to Problem (81) if and only if $\bar{Y} = \bar{A}$ is a global solution to Problem (82) with $f(\bar{Y}) = 0$. The KKT condition for Problem (82) is given by

$$\mathcal{J}_Y f(Y) - Z = 0, \quad Y \in \mathbb{SR}_+^{n \times n}, \quad Z \in \mathbb{SR}_+^{n \times n}, \quad \langle Y, Z \rangle = 0. \quad (83)$$

Let

$$F(Y) := \mathcal{J}_Y f(Y) = \frac{1}{2} (K^T(KY - B) + (KY - B)^T K).$$

Then, the KKT condition (83) becomes the following MCP

$$Y \in \mathbb{SR}_+^{n \times n}, \quad G(Y) \in \mathbb{SR}_+^{n \times n}, \quad \langle Y, F(Y) \rangle = 0. \quad (84)$$

By using the well-known Fischer function ω defined in (35), solving the MCP (84) is equivalent to the solution of the following nonsmooth equation

$$\Phi(Y) = 0, \quad (85)$$

where $\Phi_{ij}(Y) := \omega(Y_{ij}, F_{ij}(Y))$ for $i, j = 1, \dots, n$. Also, define the merit function $\phi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ by

$$\phi(Y) := \frac{1}{2} \|\Phi(Y)\|^2. \quad (86)$$

Thus, based on the functions Ψ defined in (85) and ψ in (86), one may use Algorithm 3.23 to Problem (82), where \bar{Y} is a solution. Therefore, we obtain a solution \bar{A} to the SNIEP with lower bounds (80) as follows

$$\bar{A} = \bar{Y} + \underline{A}.$$

Similarly, the global and quadratic convergence of Algorithm 3.23 for solving Problem (82) can be developed as in Section 3.

4.3 The Case of Prescribed Entries

4.3.1 The Nonsymmetric Case

In practice, one may hope some entries of A are fixed so that the required nonnegative matrix A satisfies a certain structure (e.g., the symmetric tridiagonal oscillatory matrix in vibrations [24]). The NIEP with prescribed entries aims to find a matrix $A \in \mathbb{R}_+^{n \times n}$ such that

$$AX = X\Lambda, \quad A_{ij} = (A_a)_{ij} \quad \forall (i, j) \in \mathcal{I}, \quad (87)$$

where $A_a \in \mathbb{R}_+^{n \times n}$ is a prescribed matrix. Notice that $A = P(A_{\mathcal{I}}) + P(A_{\mathcal{J}})$. Then, Problem (87) turns into a new NIEP, which amounts to determining a vector $A_{\mathcal{J}} \in \mathbb{R}_+^{|\mathcal{J}|}$ such that

$$P(A_{\mathcal{J}})X = X\Lambda - P((A_a)_{\mathcal{I}})X. \quad (88)$$

Let $K = X^T$ and $B = (X\Lambda - P((A_a)_{\mathcal{I}})X)^T$. Define the index sets $\tilde{\mathcal{I}} \subseteq \mathcal{N}$ and $\tilde{\mathcal{J}} = \mathcal{N} \setminus \tilde{\mathcal{I}}$ such that

$$P(M_{\tilde{\mathcal{I}}}) = P(M_{\mathcal{I}})^T \quad \text{and} \quad P(M_{\tilde{\mathcal{J}}}) = P(M_{\mathcal{J}})^T \quad \forall M \in \mathbb{R}^{n \times n}.$$

Then Problem (88) is equivalent to finding a global solution to the following optimization problem

$$\begin{cases} \min & f(Y_{\tilde{\mathcal{J}}}) := \frac{1}{2} \|KP(Y_{\tilde{\mathcal{J}}}) - B\|^2 \\ \text{s.t.} & P(Y_{\tilde{\mathcal{J}}}) \in \mathbb{R}_+^{n \times n} \end{cases} \quad (89)$$

in the sense that $\bar{A}_{\mathcal{J}}$ is a solution to Problem (88) if and only if $\bar{Y}_{\tilde{\mathcal{J}}}$ is a global solution to Problem (89) with $P(\bar{Y}_{\tilde{\mathcal{J}}}) = P(\bar{A}_{\mathcal{J}})^T$ and $f(\bar{Y}_{\tilde{\mathcal{J}}}) = 0$. The KKT condition for Problem (89) is given by

$$\mathcal{J}_{Y_{\tilde{\mathcal{J}}}} f(Y_{\tilde{\mathcal{J}}}) - Z_{\tilde{\mathcal{J}}} = 0, \quad Y_{\tilde{\mathcal{J}}} \geq 0, \quad Z_{\tilde{\mathcal{J}}} \geq 0, \quad \langle Y_{\tilde{\mathcal{J}}}, Z_{\tilde{\mathcal{J}}} \rangle = 0, \quad (90)$$

Let

$$F(Y_{\tilde{\mathcal{J}}}) := \mathcal{J}_{Y_{\tilde{\mathcal{J}}}} f(Y_{\tilde{\mathcal{J}}}) = (K^T(KY - B))_{\tilde{\mathcal{J}}}.$$

Then the KKT condition (90) turns into the following MCP

$$Y_{\tilde{\mathcal{J}}} \geq 0, \quad F(Y_{\tilde{\mathcal{J}}}) \geq 0, \quad \langle Y_{\tilde{\mathcal{J}}}, F(Y_{\tilde{\mathcal{J}}}) \rangle = 0. \quad (91)$$

By using the well-known Fischer function ω defined in (35), solving the MCP (91) is equivalent to the solution of the following nonsmooth equation

$$P(\Phi(Y_{\tilde{\mathcal{J}}})) = 0, \quad (92)$$

where $\Phi_j(Y_{\tilde{\mathcal{J}}}) := \omega((Y_{\tilde{\mathcal{J}}})_j, F_j(Y_{\tilde{\mathcal{J}}}))$ for $j = 1, \dots, |\tilde{\mathcal{J}}|$. Also, define the merit function $\phi : \mathbb{R}^{|\tilde{\mathcal{J}}|} \rightarrow \mathbb{R}$ by

$$\phi(Y_{\tilde{\mathcal{J}}}) := \frac{1}{2} \|P(\Phi(Y_{\tilde{\mathcal{J}}}))\|^2. \quad (93)$$

Hence, based on the functions $P(\Phi(\cdot))$ defined in (92) and ϕ in (93), one may solve Problem (89) by Algorithm 3.23, where $P(\bar{Y}_{\tilde{\mathcal{J}}})$ is a solution. Therefore, we get a solution \bar{A} to the NIEP with prescribed entries (87) as follows

$$\bar{A} = P(\bar{Y}_{\tilde{\mathcal{J}}})^T + P((A_a)_{\mathcal{I}}).$$

As in Section 3, we can establish the global and quadratic convergence of Algorithm 3.23 for solving Problem (89).

4.3.2 The Symmetric Case

The SNIEP with prescribed entries is to find a matrix $A \in \mathbb{SR}_+^{n \times n}$ such that

$$AX = X\Lambda, \quad A_{ij} = (A_a)_{ij} \quad \forall (i, j) \in \mathcal{I}, \quad (94)$$

where $A_a \in \mathbb{SR}_+^{n \times n}$ is a prescribed matrix. Define $\hat{\mathcal{I}} := \{(i, j) \mid (i, j) \in \mathcal{I}, i \leq j\}$ and $\hat{\mathcal{J}} := \{(i, j) \mid (i, j) \in \mathcal{J}, i \leq j\}$ and for a matrix $M \in \mathbb{SR}^{n \times n}$, let $M_{\hat{\mathcal{I}}}$ be the column vector of the elements M_{ij} for all $(i, j) \in \hat{\mathcal{I}}$. and define the linear operator $\hat{P} : \mathbb{R}^{|\hat{\mathcal{I}}|} \rightarrow \mathbb{SR}^{n \times n}$ by

$$\hat{P}_{ij}(M_{\hat{\mathcal{I}}}) := \begin{cases} M_{ij}, & \text{if } (i, j) \in \hat{\mathcal{I}} \text{ or } (j, i) \in \hat{\mathcal{I}}, \\ 0, & \text{otherwise.} \end{cases}$$

Thus we have $A = \hat{P}(A_{\hat{\mathcal{I}}}) + \hat{P}(A_{\hat{\mathcal{J}}})$. Therefore, Problem (94) becomes a new SNIEP, which aims to determine a vector $A_{\hat{\mathcal{J}}} \in \mathbb{R}_+^{|\hat{\mathcal{J}}|}$ such that

$$\hat{P}(A_{\hat{\mathcal{J}}})X = X\Lambda - \hat{P}((A_a)_{\hat{\mathcal{I}}})X. \quad (95)$$

Let $K = X^T$ and $B = (X\Lambda - \hat{P}((A_a)_{\hat{\mathcal{I}}})X)^T$. Then, Problem (95) is reduced to finding a global solution to the following optimization problem

$$\begin{cases} \min & f(Y_{\hat{\mathcal{J}}}) := \frac{1}{2} \|K\hat{P}(Y_{\hat{\mathcal{J}}}) - B\|^2 \\ \text{s.t.} & Y_{\hat{\mathcal{J}}} \geq 0 \end{cases} \quad (96)$$

in the sense that $\bar{A}_{\hat{\mathcal{J}}}$ is a solution to Problem (95) if and only if $\bar{Y}_{\hat{\mathcal{J}}} = \bar{A}_{\hat{\mathcal{J}}}$ is a global solution to Problem (96) with $\hat{P}(\bar{Y}_{\hat{\mathcal{J}}}) = \hat{P}(\bar{A}_{\hat{\mathcal{J}}})$ and $g(\bar{Y}_{\hat{\mathcal{J}}}) = 0$. The KKT condition for Problem (96) is given by

$$\mathcal{J}_{Y_{\hat{\mathcal{J}}}} f(Y_{\hat{\mathcal{J}}}) - Z_{\hat{\mathcal{J}}} = 0, \quad Y_{\hat{\mathcal{J}}} \geq 0, \quad Z_{\hat{\mathcal{J}}} \geq 0, \quad \langle Y_{\hat{\mathcal{J}}}, Z_{\hat{\mathcal{J}}} \rangle = 0, \quad (97)$$

Let

$$F(Y_{\hat{\mathcal{J}}}) := \mathcal{J}_{Y_{\hat{\mathcal{J}}}} f(Y_{\hat{\mathcal{J}}}) = \frac{1}{2} (K^T (K \hat{P}(Y_{\hat{\mathcal{J}}}) - B) + (K \hat{P}(Y_{\hat{\mathcal{J}}}) - B)^T K)_{\hat{\mathcal{J}}}.$$

Then, the KKT condition (97) is reduced to the following MCP

$$Y_{\hat{\mathcal{J}}} \geq 0, \quad F(Y_{\hat{\mathcal{J}}}) \geq 0, \quad \langle Y_{\hat{\mathcal{J}}}, F(Y_{\hat{\mathcal{J}}}) \rangle = 0. \quad (98)$$

By using the well-known Fischer function ω defined in (35), solving the MCP (98) is equivalent to the solution of the following nonsmooth equation

$$\hat{P}(\Phi(Y_{\hat{\mathcal{J}}})) = 0, \quad (99)$$

where $\Phi_j(Y_{\hat{\mathcal{J}}}) := \omega((Y_{\hat{\mathcal{J}}})_j, F_j(Y_{\hat{\mathcal{J}}}))$ for $j = 1, \dots, |\hat{\mathcal{J}}|$. Also, define the merit function $\phi : \mathbb{R}^{|\hat{\mathcal{J}}|} \rightarrow \mathbb{R}$ by

$$\phi(Y_{\hat{\mathcal{J}}}) := \frac{1}{2} \|\hat{P}(\Phi(Y_{\hat{\mathcal{J}}}))\|^2. \quad (100)$$

Thus, based on the functions $\hat{P}(\Phi(\cdot))$ defined in (99) and ϕ in (100), one may use Algorithm 3.23 to Problem (96), where \bar{Y} is a solution. Therefore, we get a solution \bar{A} to the SNIEP with prescribed entries (94) as follows

$$\bar{A} = \hat{P}(\bar{Y}_{\hat{\mathcal{J}}}) + \hat{P}((A_a)_{\hat{\mathcal{J}}}).$$

The global and quadratic convergence of Algorithm 3.23 for solving Problem (96) can be derived as in Section 3.

5 Numerical Tests

In this section, we report the numerical performance of Algorithm 3.23 for solving the NIEP. All the numerical tests were done using `MATLAB 7.10` on a personal computer Intel(R) Core(TM)2 Duo of 1.80 GHz CPU. As in [13], we choose a matrix Z as

$$Z_{ij} = \begin{cases} 1, & \text{if } Y_{ij}^k = 0 = F_{ij}(Y^k), \\ 0, & \text{otherwise,} \end{cases} \quad \forall i, j = 1, \dots, n,$$

for identifying an element in $\partial_B \Phi(Y^k)$ (see (42)). In our numerical experiments, we choose the starting point $Y^0 = 0$ and the stopping criterion for Algorithm 3.23 is set to be $\phi(Y^k) \leq 10^{-20}$. The parameter δ and the function θ in (46) were set as $\delta = 0.05$ and $\theta(t) = 0.1 \min\{1, t\}$. Also, we set the other parameters used in our algorithms as $\eta = 10^{-5}$, $\rho = 0.5$, and $\sigma = 10^{-4}$. For demonstration purpose, the linear system (47) was solved by the TFQMR method [20] (we set the largest number of iterations in TFQMR as $\max(100, pn)$). Of course, one may solve (47) by using other iterative methods such as the QMR method [21], the BICGSTABL method [46], the GMRES method [39], and the CGS method [42].

Example 5.1 We focus on the NIEP in Example 2.4.

By applying Algorithm 3.23 to Example 5.1, we get the following physical solution

$$\bar{A} = \begin{bmatrix} 0.8876 & 0.1437 & 0.8302 & 0.7295 & 0.4711 & 0.4780 \\ 0.5587 & 1.0091 & 0.7787 & 0.4828 & 0.4563 & 0.7778 \\ 0.9444 & 0.4134 & 0.9486 & 0.7746 & 0.5343 & 0.6240 \\ 0.5405 & 0.3926 & 0.5768 & 0.9097 & 0.7712 & 0.8098 \\ 0.3488 & 0.4854 & 0.4339 & 0.7298 & 0.6675 & 0.7527 \\ 0.5193 & 0.7965 & 0.6736 & 0.8095 & 0.7427 & 0.9366 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 4.6 \times 10^{-16}$. The convergence results are shown in Table 1.

Example 5.2 We focus on the NIEP with prescribed entries. Let $n = 5$ and \hat{A} is a random $n \times n$ nonnegative matrix given by

$$\hat{A} = \begin{bmatrix} 0.6452 & 0.3932 & 0.5707 & 0.5642 & 0.0327 \\ 0.4013 & 0.8016 & 0.5690 & 0.8279 & 0.2570 \\ 0.3559 & 0.6667 & 0.8872 & 0.5908 & 0.5805 \\ 0.2526 & 0.7224 & 0.9677 & 0.2902 & 0.9604 \\ 0.8972 & 0.9120 & 0.1895 & 0.9093 & 0.8930 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\lambda_1 = 3.0422$, $\lambda_2 = -0.3150$, $\lambda_{3,4} = 0.2801 \pm 0.3442\sqrt{-1}$, $\lambda_5 = 0.2298$. Suppose that $P((A_a)_{\mathcal{I}})$ is given by

$$P((A_a)_{\mathcal{I}}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{0.9677} & 0 & \mathbf{0.9604} \\ 0 & \mathbf{0.9120} & 0 & \mathbf{0.9093} & 0 \end{bmatrix}.$$

We choose the three eigenvalues $\{\lambda_1 = 3.0422, \lambda_{3,4} = 0.2801 \pm 0.3442\sqrt{-1}\}$ and associated eigenvectors as prescribed eigendata.

By using Algorithm 3.23 to Example 5.2, we obtain the following physical solution

$$\bar{A} = \begin{bmatrix} 0.4245 & 0.6421 & 0.5406 & 0.3647 & 0.1687 \\ 0.3488 & 0.7562 & 0.7646 & 0.6539 & 0.3112 \\ 0.3877 & 0.7051 & 0.7475 & 0.7095 & 0.5453 \\ 0.2526 & 0.7224 & \mathbf{0.9677} & 0.2902 & \mathbf{0.9604} \\ 0.8972 & \mathbf{0.9120} & 0.1895 & \mathbf{0.9093} & 0.8930 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 1.9 \times 10^{-10}$. The convergence results are displayed in Table 1.

Example 5.3 We focus on the NIEP with prescribed entries in Example 2.5, where the off-tridiagonal entries of \hat{A} are seen as prescribed entries.

By applying Algorithm 3.23 to Example 5.3, we find the following physical solution

$$\bar{A} = \begin{bmatrix} 4.7270 & 0.2055 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 2.4036 & 2.1790 & 0.5575 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1.4595 & 4.1267 & 2.0226 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1.7515 & 1.7427 & 1.7361 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 0.7977 & 3.8773 & 1.4302 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0.2471 & 4.1316 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 5.3 \times 10^{-12}$. The convergence results are listed in Table 1.

Example 5.4 We focus on the SNIEP. Let $n = 6$ and \hat{A} is a random n -by- n symmetric non-negative matrix given by

$$\hat{A} = \begin{bmatrix} 0.8270 & 0.4317 & 0.9324 & 0.6496 & 0.3172 & 0.1868 \\ 0.4317 & 1.0324 & 0.6288 & 0.3769 & 0.6401 & 0.6777 \\ 0.9324 & 0.6288 & 0.9698 & 0.5905 & 0.8683 & 0.7188 \\ 0.6496 & 0.3769 & 0.5905 & 0.9965 & 0.8499 & 0.7062 \\ 0.3172 & 0.6401 & 0.8683 & 0.8499 & 0.7110 & 0.7151 \\ 0.1868 & 0.6777 & 0.7188 & 0.7062 & 0.7151 & 0.8055 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\{-0.4176, 0.0252, 0.2241, 0.6471, 0.8334, 4.0301\}$. We use the last three eigenvalues $\{\lambda_i\}_{i=4}^6$ and associated eigenvectors $\{x_i\}_{i=4}^6$ as prescribed eigendata.

By using Algorithm 3.23 to Example 5.4, we have the following physical solution

$$\bar{A} = \begin{bmatrix} 0.9160 & 0.3769 & 0.8516 & 0.5469 & 0.4294 & 0.2504 \\ 0.3769 & 0.9857 & 0.7284 & 0.3461 & 0.6304 & 0.6933 \\ 0.8516 & 0.7284 & 1.0116 & 0.7413 & 0.7330 & 0.6249 \\ 0.5469 & 0.3461 & 0.7413 & 1.0039 & 0.7950 & 0.6936 \\ 0.4294 & 0.6304 & 0.7330 & 0.7950 & 0.7875 & 0.7678 \\ 0.2504 & 0.6933 & 0.6249 & 0.6936 & 0.7678 & 0.8058 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 5.3 \times 10^{-16}$. The convergence results are reported in Table 1.

Example 5.5 We focus on the SNIEP with prescribed entries. Let $n = 5$ and \hat{A} is a random $n \times n$ symmetric nonnegative matrix given by

$$\hat{A} = \begin{bmatrix} 0.9512 & 0.4323 & 0.3918 & 0.7250 & 0.8987 \\ 0.4323 & 0.9984 & 0.8935 & 0.2227 & 0.2427 \\ 0.3918 & 0.8935 & 0.2743 & 0.7654 & 0.5770 \\ 0.7250 & 0.2227 & 0.7654 & 0.4325 & 0.0912 \\ 0.8987 & 0.2427 & 0.5770 & 0.0912 & 0.7217 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\{-0.8234, -0.0078, 0.4695, 0.9190, 2.8208\}$. Suppose that

$P((A_a)_I)$ is given by

$$P((A_a)_I) = \begin{bmatrix} \mathbf{0.9512} & 0 & 0 & 0 & \mathbf{0.8987} \\ 0 & \mathbf{0.9984} & \mathbf{0.8935} & 0 & 0 \\ 0 & \mathbf{0.8935} & 0 & \mathbf{0.7654} & 0 \\ 0 & 0 & \mathbf{0.7654} & 0 & 0 \\ \mathbf{0.8987} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We choose the last two eigenvalues $\{\lambda_i\}_{i=4}^5$ and associated eigenvectors $\{x_i\}_{i=4}^5$ as prescribed eigendata.

By applying Algorithm 3.23 to Example 5.5, we find the following physical solution

$$\bar{A} = \begin{bmatrix} \mathbf{0.9512} & 0.3488 & 0.5927 & 0.5822 & \mathbf{0.8987} \\ 0.3488 & \mathbf{0.9984} & \mathbf{0.8935} & 0.2439 & 0.3315 \\ 0.5927 & \mathbf{0.8935} & 0.2491 & \mathbf{0.7654} & 0.3460 \\ 0.5822 & 0.2439 & \mathbf{0.7654} & 0.4019 & 0.2784 \\ \mathbf{0.8987} & 0.3315 & 0.3460 & 0.2784 & 0.7096 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 1.2 \times 10^{-10}$. The convergence results are presented in Table 1.

Example 5.6 We focus on the SNIEP with prescribed entries, which comes from the inverse problem in vibrations [24]. Let \hat{A} be a symmetric tridiagonal oscillatory matrix of order n with $n = 6$, i.e.,

$$\hat{A} = \begin{bmatrix} 4.7270 & 0.8246 & 0 & 0 & 0 & 0 \\ 0.8246 & 4.4522 & 1.1618 & 0 & 0 & 0 \\ 0 & 1.1618 & 4.9387 & 1.1349 & 0 & 0 \\ 0 & 0 & 1.1349 & 4.2360 & 1.1497 & 0 \\ 0 & 0 & 0 & 1.1497 & 4.0277 & 0.6471 \\ 0 & 0 & 0 & 0 & 0.6471 & 4.1316 \end{bmatrix}.$$

The eigenvalues of \hat{A} are given by $\{2.4625, 3.2861, 4.0556, 4.7689, 5.4343, 6.5059\}$. We use the last three eigenvalues $\{\lambda_i\}_{i=4}^6$ and associated eigenvectors $\{x_i\}_{i=4}^6$ as prescribed eigendata. The off-tridiagonal entries of \hat{A} are seen as prescribed entries.

By using Algorithm 3.23 to Example 5.6, we obtain the following physical solution

$$\bar{A} = \begin{bmatrix} 4.7270 & 0.8246 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0.8246 & 4.4522 & 1.1618 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1.1618 & 4.9387 & 1.1349 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1.1349 & 4.2360 & 1.1497 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1.1497 & 4.0277 & 0.6471 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0.6471 & 4.1316 \end{bmatrix}$$

with the error $\|\bar{A}X - \Lambda X\| < 3.7 \times 10^{-13}$. The convergence results are shown in Table 1.

From Table 1, we observe that the proposed algorithm converges quadratically.

To further illustrate the effectiveness of our method, we consider the following large-scale NIEPs.

Table 1: Convergence results for Example 5.1–5.4

	Ex. 5.1	Ex. 5.2	Ex. 5.3	Ex. 5.4	Ex. 5.5	Ex. 5.6
IT.	$\phi(Y^k)$	$\phi(Y^k)$	$\phi(Y^k)$	$\phi(Y^k)$	$\phi(Y^k)$	$\phi(Y^k)$
0	3.2×10^1	1.0×10^1	9.7×10^1	3.5×10^1	3.9×10^0	1.8×10^2
1	1.5×10^0	5.4×10^{-1}	5.2×10^0	1.6×10^0	2.2×10^{-1}	8.8×10^0
2	8.6×10^{-2}	2.7×10^{-2}	3.4×10^{-1}	8.9×10^{-2}	9.3×10^{-3}	5.2×10^{-1}
3	2.9×10^{-4}	6.6×10^{-5}	3.4×10^{-3}	2.9×10^{-4}	1.8×10^{-5}	7.1×10^{-3}
4	2.5×10^{-9}	3.1×10^{-10}	1.4×10^{-6}	2.4×10^{-9}	5.4×10^{-11}	3.5×10^{-6}
5	1.9×10^{-19}	7.1×10^{-21}	1.6×10^{-13}	1.7×10^{-19}	5.7×10^{-22}	1.6×10^{-13}
6	9.9×10^{-32}		6.8×10^{-24}	8.9×10^{-32}		4.0×10^{-27}
cputime	0.4 s	0.2 s	0.3 s	0.2 s	0.2 s	0.3 s

Example 5.7 We consider the NIEP for varying n and p . Let \hat{A} be a random $n \times n$ nonnegative matrix with each entry generated from the uniform distribution on the interval $[0, 10]$. We choose the first p eigenvalues of \hat{A} with largest absolute values and associated eigenvectors as prescribed eigendata. We report our numerical results for (a) $p = 20$ and $n = 100, 200, 500, 1000, 1500, 2000$ and (b) $p = 10, 20, 50, 100, 150, 200$ and $n = 500$.

Example 5.8 This example is concerned with the NIEP with prescribed entries for different n and p . Let \hat{A} be a random $n \times n$ nonnegative matrix with entries $\hat{A}_{ij} \in [0, 1]$ for all $i, j = 1 \dots, n$. Let the entries of \hat{A} , whose values are equal to or greater than 0.85, be the prescribed entries. We choose the first p eigenvalues of \hat{A} with largest absolute values and associated eigenvectors as prescribed eigendata. We report our numerical results for (a) $p = 20$ and $n = 100, 200, 500, 1000, 1500, 2000$ and (b) $p = 10, 20, 50, 80, 100$ and $n = 500$.

Example 5.9 We focus on the SNIEP for varying n and p . Let \hat{A} be a random $n \times n$ symmetric nonnegative matrix with each entry generated from the uniform distribution on the interval $[0, 10]$. We choose the first largest p eigenvalues of \hat{A} and associated eigenvectors as prescribed eigendata. We report our numerical results for (a) $p = 20$ and $n = 100, 200, 500, 1000, 1500, 2000$ and (b) $p = 10, 20, 50, 100, 150, 200$ and $n = 500$.

Example 5.10 We consider the SNIEP with prescribed entries for different n and p . Let \hat{A} be a random $n \times n$ symmetric nonnegative matrix with entries $\hat{A}_{ij} \in [0, 1]$ for all $i, j = 1 \dots, n$. Let the entries of \hat{A} , whose values are equal to or greater than 0.85, be the prescribed entries. We choose the first largest p eigenvalues of \hat{A} and associated eigenvectors as prescribed eigendata. We report our numerical results for (a) $p = 20$ and $n = 100, 200, 500, 1000, 1500, 2000$ and (b) $p = 10, 20, 50, 80, 100$ and $n = 500$.

The numerical results for Examples 5.7–5.10 are reported in Tables 2–5, where IT., NF., Res0., Res*, and Err. stand for the number of iterations, the number of function evaluations, the value of $\phi(\cdot)$ at the initial point Y^0 and at the final iterate of our algorithm, and the error $\|\bar{A}X - X\Lambda\|$ at the final iterate of our algorithm, respectively. Tables 2–5 put in clear evidence that the proposed algorithm is very efficient for solving large-scale NIEPs and SNIEPs.

Indeed, the proposed algorithm converges to the desired accuracy with only a small number of iterations for all cases and the quadratic convergence is also observed. The latter agrees with our theoretical results. As a final remark, we stress that our technique is quite robust since a good numerical behavior is observed independently on the problem parameters p and n .

Table 2: Numerical results for Example 5.7

$p = 20$						
n	cputime	IT.	NF.	Res0.	Res*.	Err.
100	1 s	8	9	5.1×10^5	5.6×10^{-27}	1.1×10^{-13}
200	3 s	9	10	2.0×10^6	4.6×10^{-26}	3.1×10^{-13}
500	20 s	10	11	1.3×10^7	1.9×10^{-23}	8.6×10^{-12}
1,000	46 s	8	9	5.0×10^7	1.9×10^{-23}	8.0×10^{-12}
1,500	1 m 26 s	8	9	1.1×10^8	1.3×10^{-22}	1.8×10^{-11}
2,000	4 m 05 s	10	11	2.0×10^8	1.6×10^{-24}	1.8×10^{-12}
$n = 500$						
p	cputime	IT.	NF.	Res0.	Res*.	Err.
10	12 s	8	9	1.3×10^7	1.6×10^{-25}	5.5×10^{-13}
20	29 s	11	12	1.3×10^7	1.6×10^{-25}	5.7×10^{-13}
50	32 s	13	18	1.3×10^7	1.8×10^{-24}	1.9×10^{-12}
100	2 m 08 s	16	17	1.3×10^7	5.6×10^{-25}	1.0×10^{-12}
150	3 m 58 s	20	21	1.3×10^7	3.2×10^{-22}	4.9×10^{-11}
200	37 m 19 s	39	64	1.4×10^7	6.1×10^{-24}	6.4×10^{-12}

References

- [1] R. B. Bapat and T. E. S. Raghavan, *Nonnegative Matrices and Applications*, Encyclopedia of Mathematics and its Applications 64, Cambridge University Press, 1997.
- [2] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, 9, SIAM, Philadelphia, 1994.
- [3] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York 2000.
- [4] B. Chen and P. T. Harker, *Smooth approximations to nonlinear complementarity problems*, SIAM J. Optim., 7 (1997), pp. 403–420.
- [5] X. Chen and D. L. Liu, *Isospectral flow method for nonnegative inverse eigenvalue problem with prescribed structure*, J. Comput. Appl. Math., 235 (2011), pp. 3990–4002.
- [6] M. T. Chu, F. Diele, and I. Sgura, *Gradient flow method for matrix completion with prescribed eigenvalues*, Linear Algebra Appl., 379 (2004), pp. 85–112.
- [7] M. T. Chu and K. R. Driessel, *Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations*, SIAM J. Math. Anal., 22 (1991), pp. 1372–1387.

Table 3: Numerical results for Example 5.8

$p = 20$						
n	cputime	IT.	NF.	Res0.	Res*.	Err.
100	1 s	7	8	2.4×10^3	1.1×10^{-28}	7.6×10^{-14}
200	3 s	7	8	9.1×10^3	1.8×10^{-28}	2.8×10^{-14}
500	21 s	7	8	5.6×10^4	1.2×10^{-25}	6.7×10^{-13}
1,000	1 m 23 s	7	8	2.2×10^5	5.4×10^{-21}	1.2×10^{-10}
1,500	3 m 05 s	8	9	5.0×10^5	4.3×10^{-27}	1.2×10^{-13}
2,000	5 m 04 s	8	9	8.9×10^5	6.5×10^{-27}	1.6×10^{-13}
$n = 500$						
p	cputime	IT.	NF.	Res0.	Res*.	Err.
10	14 s	7	8	5.6×10^4	8.9×10^{-27}	2.1×10^{-13}
20	26 s	8	9	5.6×10^4	1.4×10^{-22}	3.6×10^{-11}
50	51 s	8	9	5.8×10^4	7.9×10^{-28}	6.0×10^{-14}
80	1 m 19 s	8	9	6.0×10^4	6.9×10^{-26}	6.7×10^{-13}
100	1 m 29 s	8	9	6.1×10^4	6.4×10^{-21}	1.2×10^{-10}

- [8] M. T. Chu and G. H. Golub, *Structured inverse eigenvalue problems*, Acta Numer., 2002, pp. 1–71.
- [9] M. T. Chu and G. H. Golub, *Inverse Eigenvalue Problems: Theory, Algorithms, and Applications*, Oxford University Press, Oxford, 2005.
- [10] M. T. Chu and Q. Guo, *A numerical method for the inverse stochastic spectrum problem*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 1027–1039.
- [11] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.
- [12] B. N. Datta, *Numerical Methods for Linear Control Systems : Design and Analysis*, Elsevier Academic Press, 2003.
- [13] T. De Luca, F. Facchinei, and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Math. Programming, 75 (1996), pp. 407–439.
- [14] B. C. Eaves, *On the basic theorem for complementarity*, Math. Programming, 1 (1971), pp. 68–75.
- [15] P. D. Egleston, T. D. Lenker, and S. K. Narayan, *The nonnegative inverse eigenvalue problem*, Linear Algebra Appl., 379 (2004), pp. 475–490.
- [16] F. Facchinei and J. Soares, *A new merit function for nonlinear complementarity problems and a related algorithm*, SIAM J. Optim., 7 (1997), pp. 225–247.
- [17] A. Fischer, *A special Newton-type optimization method*, Optimization, 24 (1992), pp. 269–284.

Table 4: Numerical results for Example 5.9

$p = 20$						
n	cputime	IT.	NF.	Res0.	Res*.	Err.
100	0.5 s	7	8	5.3×10^5	3.9×10^{-27}	1.6×10^{-13}
200	1 s	7	8	2.1×10^6	5.1×10^{-26}	4.9×10^{-13}
500	8 s	8	9	1.3×10^7	9.4×10^{-26}	7.5×10^{-13}
1,000	34 s	9	11	5.1×10^7	2.8×10^{-25}	1.1×10^{-12}
1,500	38 s	8	9	1.1×10^8	4.1×10^{-25}	1.3×10^{-12}
2,000	1 m 10 s	8	9	2.0×10^8	5.8×10^{-25}	1.5×10^{-12}
$n = 500$						
p	cputime	IT.	NF.	Res0.	Res*.	Err.
10	7 s	8	9	1.3×10^7	9.4×10^{-26}	6.4×10^{-13}
20	7 s	8	9	1.3×10^7	9.9×10^{-26}	7.3×10^{-13}
50	8 s	8	9	1.3×10^7	1.2×10^{-25}	8.0×10^{-13}
100	8 s	8	9	1.3×10^7	1.7×10^{-25}	1.0×10^{-12}
150	9 s	8	9	1.3×10^7	2.2×10^{-25}	1.1×10^{-12}
200	10 s	8	9	1.4×10^7	2.6×10^{-25}	1.3×10^{-12}

- [18] A. Fischer, *On the local superlinear convergence of a Newton-type method for LCP under weak conditions*, Optim. Methods Softw., 6 (1995), pp. 83–107.
- [19] A. Fischer, *Solution of monotone complementarity problems with Lipschitzian functions*, Math. Programming, 76 (1997), pp. 513–532.
- [20] R. W. Freund, *A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems*, SIAM J. Sci. Comput., 14 (1993), pp. 470–482.
- [21] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.
- [22] M. I. Friswell and J. E. Mottershead, *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, 1995.
- [23] C. Geiger and C. Kanzow, *On the resolution of monotone complementarity problems*, Comput. Optim. Appl., 5 (1996), pp. 155–173.
- [24] G. M. L. Gladwell, *Inverse Problems in Vibration*, Kluwer Academic Publishers, 2004.
- [25] H. Y. Jiang and L. Q. Qi, *A new nonsmooth equations approach to nonlinear complementarity problems*, SIAM J. Control Optim., 35 (1997), pp. 178–193.
- [26] C. Kanzow, *Some noninterior continuation methods for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 851–868.
- [27] F. W. Meng, D. F. Sun, and G. Y. Zhao, *Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization*, Math. Programming, 104 (2005), pp. 561–581.

Table 5: Numerical results for Example 5.10

$p = 20$						
n	cputime	IT.	NF.	Res0.	Res*.	Err.
100	1 s	6	7	4.2×10^3	5.6×10^{-22}	6.6×10^{-11}
200	2 s	7	8	1.7×10^4	2.1×10^{-28}	3.2×10^{-14}
500	11 s	7	8	1.0×10^5	1.1×10^{-27}	7.3×10^{-14}
1,000	34 s	7	8	4.1×10^5	1.3×10^{-25}	7.6×10^{-13}
1,500	1 m 15 s	7	8	9.1×10^5	1.1×10^{-23}	7.2×10^{-12}
2,000	2 m 07 s	7	8	1.6×10^6	1.4×10^{-22}	2.5×10^{-11}
$n = 500$						
p	cputime	IT.	NF.	Res0.	Res*.	Err.
10	10 s	7	8	1.0×10^5	9.4×10^{-28}	6.6×10^{-14}
20	11 s	7	8	1.0×10^5	1.1×10^{-27}	7.6×10^{-14}
50	14 s	7	8	1.0×10^5	2.0×10^{-27}	1.0×10^{-13}
80	18 s	7	8	1.1×10^5	5.4×10^{-26}	1.4×10^{-13}
100	21 s	7	8	1.1×10^5	3.4×10^{-27}	1.4×10^{-13}

- [28] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15 (1977), pp. 957–972.
- [29] H. Minc, *Nonnegative Matrices*, John Wiley&Sons, New York, 1988.
- [30] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, 1999.
- [31] R. Orsi, *Numerical methods for solving inverse eigenvalue problems for nonnegative matrices*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 190–212.
- [32] J. S. Pang, D. F. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Math. Oper. Res., 28 (2003), pp. 39–63.
- [33] L. Q. Qi and H. Y. Jiang, *Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton methods and quasi-Newton methods for solving these equations*, Math. Oper. Res., 22 (1997), pp. 301–325.
- [34] L. Q. Qi and J. Sun, *A nonsmooth version of Newton’s method*, Math. Program., 58 (1993), pp. 353–367.
- [35] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [36] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy*, Mathematical Programming Study, 22 (1984), pp. 217–230.
- [37] R. T. Rockafellar, *Convex Analysis*, Princeton, New Jersey, 1970.

- [38] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
- [39] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.
- [40] E. Senata, *Non-negative Matrices and Markov Chains*, 2nd rev. ed., Springer-Verlag, New York, 2006.
- [41] A. Shapiro, *Sensitivity analysis of generalized equations*, J. Math. Sci., 115 (2003), pp. 2554–2565.
- [42] P. Sonneveld, *CGS: A fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 36–52.
- [43] D. F. Sun, *The strong second-order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications*, Math. Oper. Res., 31 (2006), pp. 761–776.
- [44] D. F. Sun and J. Sun, *Semismooth matrix valued functions*, Math. Oper. Res., 27 (2002), pp. 150–169.
- [45] J. G. Sun, *Backward perturbation analysis of certain characteristic subspaces*, Numer. Math., 65 (1993), pp. 357–382.
- [46] H. A. van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of the BI-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 631–644.
- [47] S. F. Xu, *An Introduction to Inverse Eigenvalue Problems*, Peking University Press and Vieweg Publishing, Beijing, 1998.
- [48] N. Yamashita and M. Fukushima, *Modified Newton methods for solving semismooth reformulations of monotone complementarity problems*, Math. Programming, 76 (1997), pp. 469–491.